Validating and Extending the Two-Moment Capital Asset Pricing Model for Financial Time Series

Serdar Neslihanoglu

A thesis submitted in fulfillment for the degree of

Doctor of Philosophy

School of Mathematics and Statistics

College of Science and Engineering

University of Glasgow

October 2014

Declaration of Authorship

I, Serdar Neslihanoglu, declare that this thesis titled, "Validating and Extending the Two-Moment Capital Asset Pricing Model for Financial Time Series" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

The work presented in Chapter 4 was presented as a poster presentation at the 4th International Disaster and Risk Conference (IDRC) in Davos, 2012, with the title *Time-varying Beta Risk of Turkish Industry Portfolios: A Comparison of GARCH and Kalman Filter Modelling Techniques*, and was jointly authored with Professor John McColl.

The part of work presented in Chapter 5 was presented as a talk at the European Conference on Data Analysis 2013 (ECDA 2013), in Luxembourg, 2013, with the title *Test of the Unconditional and Conditional Form of CAPMs in Developed and Emerging Markets.* In addition, it was presented as a talk at the 36^{th} Research Students' Conference in Probability, Statistics and Social Statistics (RSC 2013), in Lancaster, 2013, with the title *Test of the Unconditional Form of the Higher-Moment CAPMs in Developed and Emerging Markets.*

To Neslihanoglu Family

Abstract

This thesis contributes to the ongoing discussion about the financial and statistical modelling of returns on financial stock markets. It develops the asset pricing model concept which has received continuous attention for almost 50 years in the area of finance, as a method by which to identify the stochastic behaviour of financial data when making investment decisions, such as portfolio choices, and determining market risk.

The best known and most widely used asset pricing model detailed in the finance literature is the Two-Moment Capital Asset Pricing Model (CAPM) (consistent with the Linear Market Model), which was developed by Sharpe-Lintner-Mossin in the 1960s to explore systematic risk in a mean-variance framework and is the benchmark model for this thesis. However, this model has now been criticised as misleading and insufficient as a tool for characterising returns in financial stock markets. This is partly a consequence of the presence of non-normally distributed returns and non-linear relationships between asset and market returns. The inadequacies of the Two-Moment CAPM are qualified in this thesis, and the extensions are proposed that improve on both model fit and forecasting abilities.

To validate and extend the benchmark Linear Market Model, the empirical work presented in this thesis centres around three related extensions. The first extension compares the Linear Market Model's modelling and forecasting abilities with those of the time-varying Linear Market Model (consistent with the conditional Two-Moment CAPM) for 19 Turkish industry sector portfolios. Two statistical modelling techniques are compared: a class of GARCH-type models, which allow for non-constant variance in stock market returns, and state space models, which allow for the systematic covariance risk to change linearly over

time in the time-varying Linear Market Model. The state space modelling is shown to outperform the GARCH-type modelling. The second extension concentrates on comparing the performance of the Linear Market Model, with models for higher order moments, including polynomial extensions and a Generalised Additive Model (GAM). In addition, time-varying versions of the Linear Market Model and polynomial extensions, in the form of state space models, are considered. All these models are applied to 18 global markets during three different time periods: the entire period from July 2002 to July 2012, from July 2002 to just before the October 2008 financial crisis, and from after the October 2008 financial crisis to July 2012. Although the more complex unconditional models are shown to improve slightly on the Linear Market Model, the state space models again improve substantially on all the unconditional models. The final extension focuses on comparing the performance of four possible multivariate state space forms of the time-varying Linear Market Models, using data on the same 18 global markets, utilising correlations between markets. This approach is shown to improve further on the performance of the univariate state space models.

The thesis concludes by drawing together three related themes: the inappropriateness of the Linear Market Model, the extent to which multivariate modelling improves the univariate market model and the state of the world's stock markets.

Acknowledgements

I would first like to sincerely thank both of my supervisors, Professor John McColl and Dr. Duncan Lee. They guided me in the fields of statistics and finance, offered an untold amount of assistance, and had always been patient with me during my research at University of Glasgow since 2011. Many thanks are due to all of the members of the School of Mathematics and Statistics who provided me inspiration. Also, I would like to thank all my colleagues who have provided comments about my research since 2011.

I would like to thank above all my dad (Huseyin), my mom (Saniye), my brother (Musa) and his wife (Ismet), my lovely niece (Ada) and the rest of my family members. Without their whole-hearted love and support, I would not have been able to write this thesis. Finally, I would like to express my gratitude to the Republic of Turkey which provided me with financial support during my research at the University of Glasgow.

Contents

Abstract **iv**

Acknowledgements **vi**

List of Tables **xi**

List of Figures **xiii**

1 Introduction **1**

2 Financial Methodology **7**

2.1 Two-Moment Capital Asset Pricing Model 7

2.1.1 Derivation of Two-Moment CAPM from Portfolio Theory 8

2.1.2 Data Generating Process of Linear Market Model 12

2.1.3 Derivation of Two-Moment CAPM using Utility Function 14

2.2 Problems with the Two-Moment Capital Asset Pricing Model 20

2.3 Higher-Moment Capital Asset Pricing Models 22

2.3.1 Four-Moment CAPM 23

2.3.1.1 Derivation of Four-Moment CAPM using Utility Function 25

2.3.2 Three-Moment CAPM 31

2.3.3 Data Generating Processes for Higher-Moment CAPMs 31

3 Statistical Methodology **38**
3.1 Linear Models . . . 38
3.1.1 Estimation . . . 39
3.1.2 Inference . . . 40
3.2 Additive Models . . . 41
3.3 Linear Gaussian State Space Models . . . 44
3.3.1 Introduction . . . 44
3.3.2 The Kalman Filter and Smoother . . . 46
3.3.3 Estimation of Hyperparameters . . . 48
3.3.4 Kalman Filter Based Models . . . 52
3.4 Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Models . . . 53
3.4.1 Properties of GARCH-type Models . . . 54
3.4.2 Maximum Likelihood Estimation . . . 57
3.5 Model Selection and Diagnostics . . . 59
3.5.1 Model Selection . . . 59
3.5.2 Model Diagnostics . . . 61
3.5.2.1 Univariate Model Diagnostics . . . 61
3.5.2.2 Multivariate Model Diagnostics . . . 63

4 Modelling the Time-varying Systematic Covariance Risk of Turkish Industry Sector Portfolios **67**
4.1 Introduction . . . 67
4.2 Methodology . . . 70
4.2.1 Linear Market Model . . . 70
4.2.2 GARCH-type Models . . . 71
4.2.3 Kalman Filter Based Models . . . 73
4.3 Data Description . . . 74
4.4 Comparison of Models . . . 82
4.4.1 In-sample Model Fit . . . 82
4.4.2 Out-of-sample Forecasting . . . 85
4.5 Time-Varying Linear Market Model via KFMR . . . 87

4.6 Conclusion 94

5 Is the Linear Market Model appropriate for Developed and Emerging Markets? 96

5.1 Introduction 96

5.2 Methodology 99

5.2.1 Higher DGPs 99

5.2.2 Generalized Additive Model 100

5.2.3 Time-varying Higher DGPs 100

5.3 Data Description 102

5.4 Comparison of Models 110

5.4.1 Model Fit 110

5.4.2 Residual Diagnostics 114

5.4.3 Graphical Summary 118

5.4.4 Time-varying Linear Market Model 122

5.5 Comparison of Models Before and After the October 2008 Financial Crisis 129

5.5.1 Model Fit Before the October 2008 Financial Crisis 129

5.5.2 Model Fit After the October 2008 Financial Crisis 133

5.6 Conclusion 137

6 Multivariate State Space Modelling 140

6.1 Introduction 140

6.2 Methodology 144

6.3 Comparison of Models 148

6.3.1 In-sample Model Fit 148

6.3.2 Out-of-sample Forecasting 151

6.3.3 Developed markets without Japan 153

6.4 Best Forecasting Model 156

6.5 Extension of Best Forecasting Model 161

6.6 Conclusion 162

7 Conclusion and Further Work **165**

7.1 Summary 165

7.2 Key Themes 170

7.2.1 Inappropriateness of the Linear Market Model 170

7.2.2 Multivariate modelling improves univariate market modelling 171

7.2.3 State of the world's stock markets 172

7.3 Further Work 172

A R Code for Kalman Filter Mean Reverting Model **174**

Bibliography **183**

List of Tables

4.1 ISE industry sector classification. 75
4.2 Descriptive statistics of weekly returns. 77
4.3 MAE ($\times 10^2$) of in-sample model fit. 83
4.4 MSE ($\times 10^4$) of in-sample model fit. 84
4.5 MAE ($\times 10^2$) of out-of-sample forecasts. 86
4.6 MSE ($\times 10^4$) of out-of-sample forecasts. 87
4.7 Time-varying Linear Market Model hyperparameter estimates (standard errors) via KFMR. 88
4.8 Time-varying Linear Market Model state parameter estimates (standard errors) via KFMR. 89
4.9 Diagnostic test statistics for KFMR. 93

5.1 Country stock market and regional classification. 103
5.2 Descriptive statistics of weekly returns. 104
5.3 *AIC* values for all models & markets. 110
5.4 *BIC* values for all models & markets. 111
5.5 *Adjusted* R^2 values for all models & markets. 113
5.6 Normality test statistics for all models & markets. 115
5.7 Autocorrelation test statistics for all models & markets. 116
5.8 Heteroskedasticity test statistics for all models & markets. 117
5.9 Time-varying Linear Market Model hyperparameter estimates (standard errors) via KFMR. 123
5.10 Time-varying Linear Market Model state parameter estimates (standard errors) via KFMR. 124
5.11 *AIC* values for all models & markets before October 2008. 129

5.12 *BIC* values for all models & markets before October 2008. 130
5.13 *Adjusted* R^2 values for all models & markets before October 2008. 131
5.14 *AIC* values for all models & markets after October 2008. 133
5.15 *BIC* values for all models & markets after October 2008. 134
5.16 *Adjusted* R^2 values for all models & markets after October 2008. . 135

6.1 Correlation matrix between countries' stock markets. 143
6.2 MAE ($\times 10^2$) of in-sample model fit. 149
6.3 MSE ($\times 10^4$) of in-sample model fit. 150
6.4 MAE ($\times 10^2$) of out-of-sample forecasts. 152
6.5 MSE ($\times 10^4$) of out-of-sample forecasts. 153
6.6 MAE ($\times 10^2$) of in-sample model fit without Japan. 154
6.7 MSE ($\times 10^4$) of in-sample model fit without Japan. 154
6.8 MAE ($\times 10^2$) of out-of-sample forecasts without Japan. 155
6.9 MSE ($\times 10^4$) of out-of-sample forecasts without Japan. 155
6.10 Model C parameter estimates (standard errors) via KFMR for the developed markets. 157
6.11 Model C parameter estimates (standard errors) via KFMR for the emerging markets. 158
6.12 Univariate diagnostic test statistics for Model C via KFMR. . . . 159
6.13 Multivariate diagnostic test statistics for Model C via KFMR. . . 160

List of Figures

4.1 The time series plot of weekly returns on the ISE market and 4 industry sectors. 79
4.2 The time series plot of weekly returns on 7 industry sectors. . . . 80
4.3 The time series plot of weekly returns on 8 industry sectors. . . . 81
4.4 The estimated $\hat{\beta}_{imt} = \hat{\alpha}_{1it}$ plots of Bank, Chemical, Telecom, and Tourism industry sectors. 92

5.1 The time series plot of weekly returns on the MSCI World, 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets. 107
5.2 The time series plot of weekly returns on the MSCI World and 7 developed markets. 108
5.3 The time series plot of weekly returns on the MSCI World and 7 emerging markets. 109
5.4 The scatter plots of 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets weekly excess returns. 119
5.5 The scatter plots of 7 developed markets weekly excess returns. . 120
5.6 The scatter plots of 7 emerging markets weekly excess returns. . . 121
5.7 The estimated $\hat{\alpha}_{1it}$ plots of 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets. 126
5.8 The estimated $\hat{\alpha}_{1it}$ plots of 7 developed markets. 127
5.9 The estimated $\hat{\alpha}_{1it}$ plots of 7 emerging markets. 128

Chapter 1

Introduction

The concept of an asset pricing model has received continuous attention for almost 50 years from researchers working in the finance domain. Their aim has been to identify the stochastic behaviour of financial data in order to determine risk and to provide guidelines for financial investment decisions such as portfolio choice. In finance, there are different types of risk such as credit risk, business risk, liquidity risk and market or systematic risk. The purpose of an asset pricing model is to address systematic risk and this is the focus of this thesis.

The best known and most widely used asset pricing model is the Two-Moment Capital Asset Pricing Model (CAPM) which was introduced by Sharpe-Lintner-Mossin in the 1960s for exploring systematic risk in the mean-variance framework. This model assumes a linear relationship between the expected return on a financial asset and that over the whole market in which the asset is traded, summarised in a single parameter, the systematic covariance (*beta*) risk, which is assumed to be constant over time. The validity of this model depends on two restrictive assumptions, namely that asset returns are normally distributed and that the investor's utility function is quadratic, so that the distribution of wealth can be adequately summarised by just its mean and variance. However, empirical evidence (e.g. Kraus and Litzenberger (1976), Fang and Lai (1997), Hwang and Satchell (1999), Mergner and Bulla (2008) and Choudhry and Wu (2009)) suggests that the Two-Moment CAPM with constant systematic covariance risk may be misleading and insufficient to characterise asset returns since returns on assets are currently understood to be non-normally distributed and to be non-linearly related to market returns. The inadequacies of the Two-Moment CAPM have motivated financial researchers to explore alternative extensions to the benchmark Linear Market Model.

One extension in this vein has been to incorporate higher order moments into the Two-Moment CAPM. According to the literature, the Higher-Moment CAPMs, namely the Three-Moment and Four-Moment CAPMs, capture systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*) in the distributions of financial data. The theory of these was developed by Kraus and Litzenberger (1976), Fang and Lai (1997) and Hwang and Satchell (1999). Authors empirically investigated the necessity for more complicated models by fitting Higher order Data Generating Processes (DGPs) to financial data; namely the Quadratic Market Model and Cubic Market Model. They proposed several formulations of Higher order DGPs with the intention of successfully illustrating the link between Higher order DGPs and their equivalent Higher-Moment CAPMs. For example, Barone-Adesi (1985) proposed the Quadratic Market Model to be consistent with the Three-Moment CAPM that also captures *co-skewness*. Fang and Lai (1997) and Hwang and Satchell (1999) proposed the Cubic Market Model (consistent with the Four-Moment CAPM which captures both *co-skewness* and *co-kurtosis*) to explain time series returns for various sets of financial data. Similar work in the context of even higher moments is as yet unreported.

Another extension has been to allow the systematic covariance (*beta*) risk to change linearly over time in a Two-Moment CAPM. This extension will be referred to as the conditional Two-Moment CAPM in this thesis. In recent literature, researchers (e.g. Faff et al. (2000), Mergner and Bulla (2008) and Choudhry and Wu (2009)) extensively investigated the instability of systematic covariance risk for different countries and firms, by comparing the modelling and forecasting abilities of unconditional and conditional Two-Moment CAPMs.

The previously mentioned extensions of the Two-Moment CAPM occur in a univariate context. This does not utilise the correlation structure among returns on different assets as a part of the estimation process, but a significant correlation structure can be expected as a consequence of economic and financial integration, as was revealed by Yavas (2007). Prominent papers such as those by Gibbons et al. (1989), Mackinlay and Richardson (1991) and Hansen and Jagannathan (1997) have considered asset pricing models in a multivariate context with time-invariant systematic covariance risk. However, similar work in the conditional Two-Moment CAPM with time-varying systematic covariance risk in the form of multivariate state space model is yet to be undertaken.

This thesis addresses the modelling and forecasting of time-invariant and time-varying parameters in DGPs, as consistent with their equivalent CAPMs, by applying alternative statistical modelling techniques. The best known technique for modelling and forecasting time-invariant parameters in Higher order DGPs is the Ordinary Least Squares (OLS) method. Hastie and Tibshirani (1990) have developed an additive model that extends the linear model to include smooth functions of covariates, but applications to CAPMs in finance have not been identified from the literature.

The best known approaches for modelling and forecasting time-varying systematic covariance risk in a Linear Market Model are GARCH-type models (Engle (1982) and Bollerslev (1986)). These are based on estimating the conditional variances and covariances between returns on assets and a market portfolio.

Another modelling technique for time-varying parameters in DGPs is presented by the state space model. The powerful recursive algorithm for estimating the unobserved components in the state space model, which was proposed by Kalman (1960) and Kalman and Bucy (1961), is known as the Kalman Filter algorithm. In this thesis, the Kalman Filter algorithm plays a central role in modelling time-varying systematic covariance risk in a Two-Moment CAPM in both univariate and multivariate contexts. These approaches have already been applied to different countries and firms (e.g. Faff et al. (2000), Brooks et al. (2002) and Choudhry and Wu (2009)) in order to analyse time series in finance in a univariate context. Similar work on the time-varying systematic covariance risk in the Two-Moment CAPM in a multivariate context using the state space model via the Kalman Filter algorithm is yet to be undertaken.

This thesis focuses on closing the aforementioned gaps to validate and extend the benchmark Linear Market Model. The following three research aims will be addressed.

1. Evaluate the effectiveness of the widely used Two-Moment CAPM and its DGP, equivalent to the Linear Market Model, for modelling and forecasting returns in financial time series data.

2. Investigate whether a multivariate approach to modelling and forecasting systematic covariance (*beta*) risk, which allows for between market correlations, outperforms a univariate, one stock market at a time, approach.

3. Evaluate the ability of a variety of statistical modelling techniques to forecast recent returns in financial data across a variety of stock market conditions.

To investigate the above aims, the empirical work presented in this thesis centres around three related extensions. The first extension focuses on comparing the modelling and forecasting abilities of the Linear Market Model (consistent with the Two-Moment CAPM) and the time-varying Linear Market Model (consistent with the conditional Two-Moment CAPM) for 19 Turkish industry sector portfolios. Two statistical modelling techniques, the class of GARCH-type models, which allow for non-constant variance in stock market returns, and state space models, which allow for the systematic covariance risk to change linearly over time in the time-varying Linear Market Model, are used. The second extension concentrates on comparing the performance of the Linear Market Model, newly reformulated forms of Higher order DGPs as polynomial extensions (consistent with their equivalent Higher-Moment CAPMs), the Generalised Additive Model (GAM) and the time-varying versions of the Linear Market Model and polynomial extensions (consistent with their equivalent conditional Higher-Moment CAPMs), in the form of state space models, for modelling market returns in 18 global markets. The final extension focuses on comparing the performance of four possible multivariate state space forms of time-varying Linear Market Models, using data on the same 18 global markets, utilising correlations between markets.

The remainder of this introduction describes the individual chapters of this thesis in greater detail.

Chapter 2 reviews the financial models used in this thesis. These include higher order DGPs with equivalent Higher-Moment CAPMs. The Two-Moment CAPM, which is the benchmark model, not only for this thesis but also in the literature overall, and which is consistent with the Linear Market Model, is outlined. Then, the problems related to the assumptions of the Two-Moment CAPM are discussed. After this, the Higher-Moment CAPM, which extends the Two-Moment CAPM by including the Third (skewness) and Fourth (kurtosis) central moments, is outlined. Moreover, newly formulated forms of Higher order DGPs, with their equivalent Higher-Moment CAPMs, which form the basis of the model comparisons in Chapter 5, are defined.

Chapter 3 reviews the statistical methodology which will be applied throughout the thesis, focusing on both time-invariant and time-varying coefficient models. It includes a brief review of linear models and additive models for time-invariant coefficients. A large proportion of this chapter outlines time-varying coefficient models, including a state space model, as well as the Kalman Filter and Smoother. In addition, a brief review of GARCH-type models is included. The chapter concludes by discussing appropriate model selection techniques and diagnostics.

Chapter 4 compares the modelling and forecasting abilities of the Linear Market Model and time-varying Linear Market Model. GARCH-type models, such as GARCH and GJR-GARCH with normal and t (which captures heavy tails on returns) conditional distributions, and Kalman Filter based approaches, such as random coefficients (KFRC), random walk (KFRW) and mean reverting (KFMR), are utilized. Using weekly data, a comparison is generated for 19 Turkish industry sector portfolios from the period 1 August 2002 to 16 February 2012. In all cases the Istanbul Stock Exchange (ISE) All-Share index and the three-month Turkish Interbank Offer Rate (TRLIBOR) interest rate are used as proxies for the market portfolio and the risk-free rate, respectively.

Chapter 5 assesses the appropriateness of the Linear Market Model. Its performance is compared to six possible extensions. The first two are newly reformulated forms of Higher order DGPs, as simple polynomial extensions of the Linear Market Model, namely the Quadratic Market Model and the Cubic Market Model (allowing for systematic covariance, systematic skewness and systematic kurtosis risks). Using a GAM relaxes some of the assumptions underpinning these polynomial models. The next approach is the time-varying Linear Market Model (allowing for a time-varying systematic covariance risk) via KFMR. The last two are the time-varying extensions of the new formulated forms of the Higher order DGPs, namely the time-varying Quadratic Market Model and the time-varying Cubic Market Model (allowing for time-varying systematic covariance, time-varying systematic skewness and time-varying systematic kurtosis risks) via KFMR. Using weekly data, a comparison is generated for the 9 developed and 9 emerging markets, extending during the three time periods: the entire period from July 2002 to July 2012, from July 2002 to before the October 2008 financial crisis, and from after the October 2008 financial crisis to July 2012. In all cases, the Morgan Stanley Capital International (MSCI) World Index and the three-month US dollar London Interbank Offered Rate (LIBOR) interest rate serve as the proxies for the market portfolio and the risk-free rate, respectively.

Chapter 6 models and forecasts the time-varying systematic covariance risks for the same 18 global markets based on multivariate state space forms of the time-varying Linear Market Model, using a KFMR model. To determine whether this outperforms a univariate approach, in-sample modelling and out-of-sample forecasting performance is considered. Four possible multivariate state space model formulations are considered.

Finally, Chapter 7 discusses the main conclusions of this thesis and the key themes which should affect future research related to this topic, while also providing an outline of some possible extensions and further work.

Chapter 2

Financial Methodology

This chapter describes the financial models used in this thesis, which includes the Capital Asset Pricing Models (CAPM, Sharpe-Lintner-Mossin (1960s), Kraus and Litzenberger (1976), Fang and Lai (1997) and Hwang and Satchell (1999)). In section 2.1 we derive the Two-Moment CAPM, which is the benchmark model both in this thesis and in the financial literature. In section 2.2 we discuss the problems with the Two-Moment CAPM. In section 2.3 we derive the Higher-Moment CAPM, which extends the Two-Moment CAPM by including the Third (skewness) and Fourth (kurtosis) central moments.

2.1 Two-Moment Capital Asset Pricing Model

The Two-Moment (also known as Traditional or Mean-Variance) Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965) and Mossin (1966) is the best known and most widely used approach for modelling financial risk. The idea of this model is to derive a theory of asset valuation in an equilibrium situation, which relates an asset expected return to the market risk. The Two-Moment CAPM, in equilibrium, can be represented as

$$E(R_i) - R_f = \beta_{im}[E(R_m) - R_f], \quad i \,\epsilon\, \{1, \ldots, N\}\,, \tag{2.1}$$

where R_i, $E(R_i)$ and R_f are the return and the expected return on asset i and the risk-free rate of return, respectively. The left hand side of (2.1) represents the expected excess return on asset i relative to a risk-free return, and is related to a product of two terms. Here, $[E(R_m) - R_f]$ is the expected excess return on the market portfolio (R_m) relative to the risk-free return, and β_{im} quantifies the risk of investment of asset i, and has been called the market risk, systematic risk, market beta or systematic covariance (*beta*). In this thesis, we use the terminology systematic covariance. This is defined as

$$\beta_{im} = \frac{Cov(R_i, R_m)}{\sigma(R_m)^2} = \frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]}, \tag{2.2}$$

where R_m, $E(R_m)$ and $\sigma(R_m)^2$ are the return, the expected return and variance on the market portfolio. $Cov(R_i, R_m)$ is the covariance between the return on asset i and on the market portfolio.

The following section is organized as follows. In section 2.1.1, we present how to derive the Two-Moment CAPM from portfolio theory. In section 2.1.2, we discuss how to estimate the systematic covariance (β_{im}) from a Data Generating Process (DGP), which in the financial literature is called the Linear Market Model. In section 2.1.3, we present how to derive the Two-Moment CAPM using a utility function. The portfolio theory approach is conceptually simpler, but the utility function approach will be extended in section 2.3 when considering Higher-Moment CAPMs.

2.1.1 Derivation of Two-Moment CAPM from Portfolio Theory

To derive the Two-Moment CAPM, we follow the approach of portfolio optimization by the mean-variance principle, which was developed by Markowitz (1952).[1] According to Markowitz (1952), the investor aims to achieve the highest possible expected return for a given variance of return.

Assuming the initial wealth, W_0 is 1, the investor wealth at the end of the

[1]The derivation of Two-Moment CAPM from portfolio theory throughout this thesis follows that of Levy and Sarnat (1994, App. 12A).

period, W_F is defined as

$$W_F = x_0(1 + R_f) + \sum_{i=1}^{N} x_i(1 + R_i), \tag{2.3}$$

where R_f is the return on the single riskless asset and R_i is the return on the i^{th} of N risky assets $i \in \{1, \ldots, N\}$. Here, x_0 and x_i are the proportions of the initial wealth invested in a riskless asset and the N risky assets, respectively, where $x_0 + \sum_{i=1}^{N} x_i = 1$.

To simplify the analysis, we assume that the rate of return of the portfolio held by an investor is defined as

$$R_p = \frac{W_F - W_0}{W_0}, \tag{2.4}$$

where W_0 is the initial and W_F is the final wealth of the investor. Assuming the initial wealth W_0 is 1, the portfolio return, R_p, can be represented by

$$R_p = \frac{W_F - W_0}{W_0} = x_0(1 + R_f) + \sum_{i=1}^{N} x_i(1 + R_i) - 1. \tag{2.5}$$

This can be further simplified to

$$\begin{aligned} R_p &= x_0 + x_0 R_f + \sum_{i=1}^{N} x_i + \sum_{i=1}^{N} x_i R_i - 1, \\ R_p &= x_0 R_f + \sum_{i=1}^{N} x_i R_i, \end{aligned}$$

using the relation

$$x_0 + \sum_{i=1}^{N} x_i = 1.$$

Here, subscript p is referring to the portfolio of the single risk free asset and N risky assets. The mean and standard deviation (*volatility*, *risk*) of the portfolio return, R_p, can be represented by

$$E(R_p) = x_0 R_f + \sum_{i=1}^{N} x_i E(R_i) = R_f + \sum_{i=1}^{N} x_i (E(R_i) - R_f), \tag{2.6}$$

which again uses the fact that $x_0 + \sum_{i=1}^{N} x_i = 1$.

$$\sigma(R_p) = \left[\sum_{i=1}^{N} x_i^2 \sigma_{ii} + \sum_{i=1}^{N} \sum_{j=1,\, j \neq i}^{N} x_i x_j \sigma_{ij} \right]^{1/2}, \tag{2.7}$$

where the risk free rate is a constant and not subject to variation. In addition,

$$\sigma_{ij} = E\left[(R_i - E(R_i)) (R_j - E(R_j)) \right],$$

the covariance between the returns of assets i and j.

Now, we will discuss that for a given $E(R_p)$, the individual investor should allocate wealth among assets so as to minimize the standard deviation (*volatility*, *risk*), $\sigma(R_p)$ of his portfolio subject to $E(R_p) = R_f + \sum_{i=1}^{N} x_i[E(R_i) - R_f]$. To determine how best to allocate wealth (i.e. choose the values of x_i) according to this criterion, a Lagrangian multiplier approach can be adopted. Define the function L as follows

$$\begin{aligned} L &= \sigma(R_p) + \lambda \left(E(R_p) - R_f - \sum_{i=1}^{N} x_i (E(R_i) - R_f) \right), \qquad (2.8) \\ L &= \left[\sum_{i=1}^{N} x_i^2 \sigma_{ii} + \sum_{i=1}^{N} \sum_{j=1,\, j \neq i}^{N} x_i x_j \sigma_{ij} \right]^{1/2} \\ &+ \lambda \left(E(R_p) - R_f - \sum_{i=1}^{N} x_i (E(R_i) - R_f) \right), \end{aligned}$$

where λ is the Lagrangian multiplier and the expression in brackets equals zero. After taking first order derivatives of the Lagrangian form in (2.8) with respect to proportions x_i and setting these derivatives equal to zero, the following equations are obtained.

$$\begin{aligned}\frac{\partial L}{\partial x_i} &= \frac{1}{2\sqrt{\sigma(R_p)^2}}\left\{2x_i\sigma_{ii} + 2\sum_{j\neq i} x_j\sigma_{ij}\right\} - \lambda(E(R_i) - R_f) = 0, \qquad (2.9)\\ &= \frac{1}{\sigma(R_p)}\left\{x_i\sigma_{ii} + \sum_{j\neq i} x_j\sigma_{ij}\right\} - \lambda(E(R_i) - R_f) = 0,\end{aligned}$$

for $i \ \epsilon \ \{1, \ldots, N\}$. Then multiplying both sides of the set of equations (2.9) by x_i gives

$$\frac{1}{\sigma(R_p)}\left\{x_i^2\sigma_{ii} + \sum_{j\neq i} x_i x_j\sigma_{ij}\right\} = \lambda x_i(E(R_i) - R_f). \qquad (2.10)$$

Summing equation (2.10) over $i \ \epsilon \ \{1, \ldots, N\}$ gives

$$\frac{1}{\sigma(R_p)}\left\{\sum_{i=1}^{N} x_i^2\sigma_{ii} + \sum_{i=1}^{N}\sum_{j\neq i} x_i x_j\sigma_{ij}\right\} = \lambda\sum_{i=1}^{N} x_i(E(R_i) - R_f), \qquad (2.11)$$

where the term in the brackets on the left hand side of equation (2.11) is the variance of the portfolio $\{\sigma(R_p)\}^2$.

Now, assuming that $x_0 = 0$, that is no wealth is assigned to the riskless asset gives $\sum_{i=1}^{N} x_i = 1$. Then, we obtain the following equation.

$$\begin{aligned}\frac{1}{\sigma(R_p)}\{\sigma(R_p)\}^2 &= \lambda\left(\sum_{i=1}^{N} x_i E(R_i) - \sum_{i=1}^{N} x_i R_f\right), \qquad (2.12)\\ \frac{1}{\sigma(R_p)}\{\sigma(R_p)\}^2 &= \lambda\left(E(R_p) - R_f\right),\\ \frac{1}{\lambda} &= \frac{E(R_p) - R_f}{\sigma(R_p)}.\end{aligned}$$

Rearranging equation (2.9) and substituting in equation (2.12) gives

$$\begin{aligned}E(R_i) - R_f &= \frac{1}{\lambda\sigma(R_p)}\left\{x_i\sigma_{ii} + \sum_{j\neq i} x_j\sigma_{ij}\right\}, \qquad (2.13)\\ &= \frac{E(R_p) - R_f}{\{\sigma(R_p)\}^2}\left\{x_i\sigma_{ii} + \sum_{j\neq i} x_j\sigma_{ij}\right\}.\end{aligned}$$

Now as $x_0 = 0$ here this special case is known as the market portfolio, that is

$$R_m = \sum_{i=1}^{N} x_i R_i, \tag{2.14}$$

so R_p is replaced by R_m. Then, the covariance between the return on the market, R_m and the return on asset i, R_i is given by

$$Cov(R_i, R_m) = \sum_{j=1}^{N} x_j Cov(R_i, R_j) = \left\{ x_i \sigma_{ii} + \sum_{j \neq i} x_j \sigma_{ij} \right\}. \tag{2.15}$$

Then, equation (2.13) can be written as

$$E(R_i) - R_f = \frac{E(R_m) - R_f}{\sigma(R_m)^2} \left\{ x_i \sigma_{ii} + \sum_{j \neq i} x_j \sigma_{ij} \right\}, \tag{2.16}$$

$$E(R_i) - R_f = \frac{E(R_m) - R_f}{\sigma(R_m)^2} Cov(R_i, R_m), \tag{2.17}$$

$$E(R_i) - R_f = \beta_{im}[E(R_m) - R_f], \tag{2.18}$$

where

$$\beta_{im} = \frac{Cov(R_i, R_m)}{\sigma(R_m)^2}, \tag{2.19}$$

which is the Two-Moment CAPM defined earlier.

2.1.2 Data Generating Process of Linear Market Model

The Linear Market Model is the most widely used statistical model in finance to estimate the risk measure systematic covariance (*beta*) in the Two-Moment CAPM. The Data Generating Process (DGP) of the Linear Market Model can be represented by

$$R_i - R_f = \kappa_i + \alpha_{1i}(R_m - R_f) + \varepsilon_i, \tag{2.20}$$

which is a simple linear regression of the response $R_i - R_f$ on the single covariate $R_m - R_f$. The errors are assumed to be independent and identically distributed with $E(\varepsilon_i)$=0 and $Var(\varepsilon_i)$=σ^2.

The Linear Market Model (equation (2.20)) is consistent with the Two-Moment CAPM (equation (2.1)) as taking expectations in (2.20) gives equation (2.1) if κ_i=0 and α_{1i}=β_{im}. However, the systematic covariance β_{im} in Two-Moment CAPM can be expressed as

$$\beta_{im} = \alpha_{1i}, \tag{2.21}$$

without the need for the assumption κ_i=0.

Proof: Start by taking expectations of the Linear Market Model, which gives

$$E(R_i - R_f) = \kappa_i + \alpha_{1i}E(R_m - R_f) + E(\varepsilon_i). \tag{2.22}$$

Subtract equation (2.22) from equation (2.20)

$$\begin{aligned} (R_i - R_f) - E(R_i - R_f) &= \alpha_{1i}\left((R_m - R_f) - E(R_m - R_f)\right) + \varepsilon_i, \quad (2.23) \\ R_i - E(R_i) &= \alpha_{1i}(R_m - E(R_m)) + \varepsilon_i. \end{aligned}$$

Multiply both sides of equation (2.23) by $R_m - E(R_m)$

$$\begin{aligned} (R_i - E(R_i))(R_m - E(R_m)) &= \alpha_{1i}(R_m - E(R_m))(R_m - E(R_m)) \quad (2.24) \\ &+ \varepsilon_i(R_m - E(R_m)). \end{aligned}$$

Take expected values both sides of equation (2.24)

$$E[(R_i - E(R_i))(R_m - E(R_m))] = \alpha_{1i}E[(R_m - E(R_m))^2]. \tag{2.25}$$

Divide both sides in equation (2.25) by the variance of the market return, $E[(R_m - E(R_m))^2]$

$$\frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} = \alpha_{1i}\frac{E[(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^2]}, \tag{2.26}$$

giving

$$\beta_{im} = \frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} = \alpha_{1i}. \tag{2.27}$$

Thus to estimate β_{im} one simply fits the Linear Market Model to financial data and obtains the estimates of $\hat{\alpha}_{1i}$. However, the Linear Market Model (2.20) only contains a single data point $R_i - R_f$, as each asset i has its own risk parameter β_{im}=α_{1i}.

Throughout this thesis, the parameters of the Linear Market Model, including β_{im}=α_{1i}, must be estimated from more data, usually time series data of returns at regular (for example, daily or monthly) intervals. Assuming that the parameters are constant over the whole time period being considered, then the excess return on asset i in period t ($R_{it}-R_{ft}$) might be assumed to be generated by the following model[2]

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1i}(R_{mt} - R_{ft}) + \varepsilon_{it}, \quad t \,\epsilon\, \{1, \ldots, T\}, \tag{2.28}$$

where R_{it}, R_{ft} and R_{mt} are the returns on asset i, the risk-free rate and the market portfolio at time t. The errors $\varepsilon_{it} \sim N(0, \sigma^2)$ so the maximum likelihood estimator of $\beta_{im} = \alpha_{1i}$ is

$$\hat{\beta}_{im} = \hat{\alpha}_{1i} = \frac{\sum_{t=1}^{T} \left[(R^*_{it} - \bar{R}^*_i)(R^*_{mt} - \bar{R}^*_m)\right]}{\sum_{t=1}^{T} \left[(R^*_{mt} - \bar{R}^*_m)^2\right]},$$

where R^*_{mt}=$R_{mt} - R_{ft}$, R^*_{it}=$R_{it} - R_{ft}$ and $\bar{R}^*_m$=$\frac{1}{T}\sum_{t=1}^{T} R^*_{mt}$, $\bar{R}^*_i$=$\frac{1}{T}\sum_{t=1}^{T} R^*_{it}$. Here, R^*_{it} and R^*_{mt} are the excess returns on asset i and the market portfolio at time t, while $\bar{R}^*_i$ and $\bar{R}^*_m$ are the expected excess returns on asset i and the market portfolio overall time points.

2.1.3 Derivation of Two-Moment CAPM using Utility Function

To lead into the derivation of Higher-Moment CAPMs, we now re-derive the Two-Moment CAPM using a utility function and expanding it in a Taylor series

[2]See details in Kraus and Litzenberger (1976) and Hung et al. (2004).

expansion.[3] In this derivation we define the individual's expected utility function for the Two-Moment CAPM and then maximize it as a Lagrangian form. Finally, we move from the individual equilibrium model to a market equilibrium model, and the Two-Moment CAPM is obtained. Consider any arbitrary utility function of the individual's wealth, $U(W_F)$, which can be approximated by an n^{th} order Taylor series expansion

$$\begin{aligned} U(W_F) &= U[E(W_F)] + U^{'}[E(W_F)][W_F - E(W_F)] \\ &+ \frac{1}{2!}U^{''}[E(W_F)][W_F - E(W_F)]^2 \\ &+ \sum_{n=3}^{\infty} \frac{1}{n!}U^n[E(W_F)][W_F - E(W_F)]^n, \end{aligned} \tag{2.29}$$

where $E(W_F)$ is the expected individual wealth at the end of period. Here, $U^{'}(.)$, $U^{''}(.)$ and $U^n(.)$ denote the first, second and n^{th} derivative of the utility function. Previously, we discussed the Markowitz's mean-variance approach, which provides that the investor aims to achieve the highest possible expected return, for a given variance of that return. Under this approach, the utility function above implicitly depends only on the mean and variance of the end of period wealth. Therefore, assuming that, the utility function is quadratic and the returns are normally distributed, allows us to write the expected utility function[4] only in terms of the mean and variance of W_F. This result allows us to derive the Two-Moment CAPM. Under this assumption the utility function is quadratic, the higher order derivatives are equal to zero ($U^n(.)$=0, $n \geq 3$). Then taking expectations of equation (2.29) gives[5]

$$E[U(W_F)] = U[E(W_F)] + \frac{1}{2!}U^{''}[E(W_F)]\sigma(W_F)^2, \tag{2.30}$$

since $E[W_F - E(W_F)]$=0 and $E[W_F - E(W_F)]^2$=$\sigma(W_F)^2$.

To simplify the analysis, it is possible to show the link between the end of period wealth, W_F (equation (2.3) and the portfolio return, R_p, (equation (2.5))

[3] The derivation of CAPMs using utility function throughout this thesis follows that of Hwang and Satchell (1999), Ranaldo and Favre (2005), Liow and Chan (2005) and Ziemann (2004).

[4] It is also known as Von Neumann-Morgenstern Utility.

[5] See details in Cuthbertson and Nitzsche (2005) and Pennacchi (2008, Chap. 4).

while assuming the initial wealth W_0 is 1. The following equations are obtained.

$$R_p = \frac{W_F - 1}{1} = x_0 R_f + \sum_{i=1}^{N} x_i R_i. \tag{2.31}$$

So we have

$$\begin{aligned} E(R_p) &= E(W_F) - 1, & (2.32) \\ \sigma(R_p) &= \sigma(W_F). & (2.33) \end{aligned}$$

Now to derive the Two-Moment CAPM, we also need to show that

$$\sigma(W_F) = \sum_{i=1}^{N} x_i \beta_{ip} \sigma(R_p), \tag{2.34}$$

where

$$\beta_{ip} = \frac{E[(R_i - E(R_i))(R_p - E(R_p))]}{E[(R_p - E(R_p))^2]}, \tag{2.35}$$

is the systematic risk measure of an asset i relative to the portfolio variance. The proof requires us to show that $\sum_{i=1}^{N} x_i \beta_{ip} = 1$.

$$\begin{aligned} \sum_{i=1}^{N} x_i \beta_{ip} &= \sum_{i=1}^{N} x_i \frac{E[(R_i - E(R_i))(R_p - E(R_p))]}{E[(R_p - E(R_p))^2]}, & (2.36) \\ &= \frac{E\left[\left(\sum_{i=1}^{N} x_i R_i - \sum_{i=1}^{N} x_i E(R_i)\right)(R_p - E(R_p))\right]}{E[(R_p - E(R_p))^2]}, \\ &= \frac{E\left[((R_p - x_0 R_f) - (E(R_p) - x_0 R_f))(R_p - E(R_p))\right]}{E\left[(R_p - E(R_p))^2\right]}, \\ &= \frac{E\left[(R_p - E(R_p))(R_p - E(R_p))\right]}{E\left[(R_p - E(R_p))^2\right]} = 1, \end{aligned}$$

using equations (2.5) and (2.6).

Now, the next step is to maximize the individual investor's expected utility of the end of period wealth, W_F subject to a budget constraint, that is

$$\max_{\{x_0,x_1,\ldots,x_N\}} E[U(W_F)], \tag{2.37}$$
$$\text{subject to} \quad x_0+\sum_{i=1}^{N} x_i = 1.$$

A Lagrangian multiplier approach can be adopted in order to maximize the individual investor's expected utility and solve the equilibrium condition in equation (2.37). Define the function L as follows

$$L = E[U(W_F)] - \lambda\left(x_0+\sum_{i=1}^{N} x_i - 1\right), \tag{2.38}$$

and note that as $E[U(W_F)]{=}U[E(W_F)]+\frac{1}{2!}U''[E(W_F)]\sigma(W_F)^2$, then we can write

$$E[U(W_F)] = f\left(E(W_F),\, \sigma(W_F)\right),$$

as a function of $(E(W_F),\, \sigma(W_F))$. Hence

$$L = f\left(E(W_F),\, \sigma(W_F)\right) - \lambda\left(x_0+\sum_{i=1}^{N} x_i - 1\right). \tag{2.39}$$

Here, λ is the Lagrangian multiplier and the expression in the brackets is equal to zero. Now, we start taking the first order total derivatives of the Lagrangian form (which is necessary and sufficient conditions for a maximum[6]) in equation (2.39) with respect to proportions, x_0 and x_i, and setting these derivatives equal to zero. Note that implicitly, these total derivations, which use the fact (from equation (2.39)) that $E[U(W_F)]$ is a function of $E(W_F)$ as a function of $x_0, x_1, \ldots, x_N$ and $\sigma(W_F)$ as a function of $x_1, \ldots, x_N$, are obtained from multivariate calculus as

[6]The first order derivative of the Lagrangian form is necessary and sufficient conditions for a maximum, because that of the second order has a negative definite Hessian matrix. (Tepmony (2010))

follows.

$$\begin{aligned}
\frac{\partial L}{\partial x_0} &= \frac{\partial E[U(W_F)]}{\partial E(W_F)}\frac{\partial E(W_F)}{\partial x_0} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\frac{\partial \sigma(W_F)}{\partial x_0} - \lambda = 0, \qquad (2.40)\\
&= \frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+R_f) - \lambda = 0,
\end{aligned}$$

using

$$\begin{aligned}
\frac{\partial E(W_F)}{\partial x_0} &= (1+R_f), \quad \text{from equation (2.3)}, \qquad (2.41)\\
\frac{\partial \sigma(W_F)}{\partial x_0} &= 0, \quad \text{from equation (2.34)}. \qquad (2.42)
\end{aligned}$$

Similarly, the first derivative of equation (2.39) with respect to x_i yields

$$\begin{aligned}
\frac{\partial L}{\partial x_i} &= \frac{\partial E[U(W_F)]}{\partial E(W_F)}\frac{\partial E(W_F)}{\partial x_i} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\frac{\partial \sigma(W_F)}{\partial x_i} - \lambda = 0, \qquad (2.43)\\
&= \frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+E(R_i)) + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\beta_{ip}\sigma(R_p) - \lambda = 0,
\end{aligned}$$

using

$$\begin{aligned}
\frac{\partial E(W_F)}{\partial x_i} &= 1+E(R_i), \quad \text{from equation (2.3)}, \qquad (2.44)\\
\frac{\partial \sigma(W_F)}{\partial x_i} &= \beta_{ip}\sigma(R_p), \quad \text{from equation (2.34)}. \qquad (2.45)
\end{aligned}$$

Rearranging equations (2.40) and (2.43) gives

$$\begin{aligned}
\frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+R_f) &= \frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+E(R_i)) + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\beta_{ip}\sigma(R_p),\\
(1+R_f) &= (1+E(R_i)) + \frac{\dfrac{\partial E[U(W_F)]}{\partial \sigma(W_F)}}{\dfrac{\partial E[U(W_F)]}{\partial E(W_F)}}\beta_{ip}\sigma(R_p), \qquad (2.46)\\
E(R_i) - R_f &= -\frac{\dfrac{\partial E[U(W_F)]}{\partial \sigma(W_F)}}{\dfrac{\partial E[U(W_F)]}{\partial E(W_F)}}\beta_{ip}\sigma(R_p),
\end{aligned}$$

where $-\frac{\partial E[U(W_F)]}{\partial \sigma(W_F)} / \frac{\partial E[U(W_F)]}{\partial E(W_F)}$ equals the investor's marginal rate of substitution (which is defined by minus the slope of the expected utility curve (indifference curve)[7]) between the expected and the standard deviation of the end of period wealth, W_F. Then, we rearrange the investor's marginal rate of substitution computed by partial differentiation. Start by taking the total differential of the expected utility function, $E[U(W_F)]{=}f(E(W_F), \sigma(W_F))$ from (2.30), which gives

$$dE[U(W_F)] = \frac{\partial E[U(W_F)]}{\partial E(W_F)} dE(W_F) + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)} d\sigma(W_F). \tag{2.47}$$

Divide both sides in equation (2.47) by $dE(W_F)$

$$\frac{dE[U(W_F)]}{dE(W_F)} = \frac{\partial E[U(W_F)]}{\partial E(W_F)} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)} \frac{d\sigma(W_F)}{dE(W_F)}. \tag{2.48}$$

Know that through any point on the expected utility curve, $\frac{dE[U(W_F)]}{dE(W_F)}{=}0$, because $E[U(W_F)]$ is constant, then, following equation is obtained.

$$\frac{dE[U(W_F)]}{dE(W_F)} = \frac{\partial E[U(W_F)]}{\partial E(W_F)} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)} \frac{d\sigma(W_F)}{dE(W_F)} = 0. \tag{2.49}$$

Rearranging equation (2.49) yields

$$\frac{dE(W_F)}{d\sigma(W_F)} = -\frac{\dfrac{\partial E[U(W_F)]}{\partial \sigma(W_F)}}{\dfrac{\partial E[U(W_F)]}{\partial E(W_F)}}. \tag{2.50}$$

Substituting (2.50) in equation (2.46) gives

$$E(R_i) - R_f = \frac{dE(W_F)}{d\sigma(W_F)} \sigma(R_p)\beta_{ip}. \tag{2.51}$$

The final step is to proceed from the individual equilibrium model to a market equilibrium model. For this case, the individual investor optimum portfolio is equivalent to the market portfolio; so, R_p is replaced by R_m. Then, equation

[7] See details in Pennacchi (2008, Chap. 2) and Hwang and Satchell (1999).

(2.51) can be written as

$$E(R_i) - R_f = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m)\beta_{im}, \tag{2.52}$$

where R_m is the return on the market portfolio, $\sigma(R_m)$ is the standard deviation of the market portfolio return. Here, β_{im} is the systematic risk measure of an asset i relative to market variance. Clearly, this equivalent to the Two-Moment CAPM (equation (2.1)) if

$$E(R_m) - R_f = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m). \tag{2.53}$$

This is true because if we apply equation (2.51) to the return on the entire market portfolio, R_m, we get

$$E(R_m) - R_f = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m)\beta_{mm}, \tag{2.54}$$

and $\beta_{mm}=\dfrac{Cov(R_m, R_m)}{\sigma^2(R_m)}=1$. Hence, the result is proved.

2.2 Problems with the Two-Moment Capital Asset Pricing Model

The most widely used asset pricing model detailed in the financial literature is the Two-Moment CAPM, a simple asset pricing theory, as discussed in section 2.1. However, this model has recently been criticised as misleading and insufficient to adequately characterise stock market returns. This may partly be a consequence of its restrictive assumptions. Therefore, possible extensions have been proposed in the literature to relax some of assumptions underlying the Two-Moment CAPM. These are discussed below.[8]

One extension of the Two-Moment CAPM is intended to answer criticism of the existence of a riskless asset. Black (1972) derived a Zero-Beta model demon-

[8]The extensions of the Two-Moment CAPM are mainly discussed and primarily followed from Ziemann (2004).

strating that the results of the Two-Moment CAPM do not in fact require the existence of a riskless asset. The next extension of the Two-Moment CAPM deals with criticism of the single time period assumption, in which investors maximize their portfolio at the end of the current period, so that there does not exist any opportunity for investors to restore their portfolios repeatedly over time. Therefore, Merton (1973) derived a multi-period model known as the Intertemporal CAPM (ICAPM). Another extension relaxes the assumption that the market portfolio is not observable. Breeden (1979) derived the Consumption CAPM (CCAPM) as an extension of the multi-period model, while using the consumption growth rate instead of market portfolio returns when explaining asset returns. In addition, Roll (1977) derived an arbitrage asset pricing model (APT) without a market portfolio return. A further extension allows for the size and value of the assets which cannot be explained by the Two-Moment CAPM while explaining asset returns. Investors assume that the size and value of assets can affect the expected return of assets, so the risk level of assets can change in terms of their size and value. Hence, Fama and French (1993) derived the Fama-French model, extending the Two-Moment CAPM to fit additional factors such as size and value factors. Since then, other authors have also added additional factors. All of these led to a better description of the asset returns. The majority of these extensions, however, lacked simple interpretations in terms of risk.

This thesis will criticise the simple assumptions of normally distributed asset returns and quadratic utility (as summarised by mean and variance) as well as linear relationships between asset and market returns. This follows the literature where, for example, Kraus and Litzenberger (1976), Fang and Lai (1997) and Hwang and Satchell (1999) derived Higher-Moment CAPMs, which extend the Two-Moment CAPM by including higher moments, as will be discussed in the section 2.3. These extensions are derived from a single factor (market portfolio return) instead of developing new factors. Hence, these models provide a simple interpretation in terms of risk, comparative to the other Two-Moment CAPM extensions, and are also easier to put into practice.

2.3 Higher-Moment Capital Asset Pricing Models

Previously, we discussed the Two-Moment CAPM depending on two restrictive assumptions, that asset returns are normally distributed and the utility function is quadratic (so it can be expressed in terms of the mean and variance of wealth only). However, literature (e.g. Kraus and Litzenberger (1976), Fang and Lai (1997) and Hwang and Satchell (1999)) suggests that the Two-Moment CAPM may be misleading and insufficient to characterize asset returns. Returns on many assets are now believed to be non-normally distributed, so they are characterized by their skewness, lack of symmetry of the return distribution around its mean, and kurtosis, relative peakness and flatness of a distribution compared with the normal distribution, as well as their mean and variance. In addition, literature (e.g. Roll (1977)) makes the criticism that a quadratic utility function suggests that the investor's risk aversion increases instead of decreases with increasing wealth.

All of these inadequacies of the Two-Moment CAPM would encourage financial research to explore further beyond this benchmark model. Hence, research frameworks incorporate higher order moments into the Two-Moment CAPM. In the literature, Higher-Moment CAPMs, namely Three-Moment and Four-Moment CAPMs include higher order moments, such as skewness and kurtosis. The Three-Moment CAPM was developed by Kraus and Litzenberger (1976). Moreover, Friend and Westerfield (1980), Barone-Adesi (1985), Lim (1989), Harvey and Siddique (2000a) and Harvey and Siddique (2000b) all identify skewness, characterizing the degree of symmetry of a return distribution around its mean, as playing an important role for asset pricing in the Three-Moment CAPM. The Four-Moment CAPM was investigated by Fang and Lai (1997), Hwang and Satchell (1999), Christie-David and Chaudhry (2001), Galagedera et al. (2002), Ranaldo and Favre (2005), Liow and Chan (2005), Jurczenko and Maillet (2006) and Javid (2009); and they also identify kurtosis, characterizing the relative peakness and flatness of a distribution compared with the normal distribution.

This section is organized as follows. In section 2.3.1-2.3.2 we derive the Higher-Moment CAPMs, namely Four- and Three-Moment CAPM. In section 2.3.3

we discuss how to estimate the systematic risk measures, systematic covariance (*beta*), systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*) in the Higher-Moment CAPMs from Data Generating Processes (DGPs), which in the literature is called the Cubic Market Model and Quadratic Market Model.

2.3.1 Four-Moment CAPM

The Four-Moment CAPM, in equilibrium, can be represented as

$$E(R_i) - R_f = c_1\beta_{im} + c_2\gamma_{im} + c_3\delta_{im}, \tag{2.55}$$

where $E(R_i)$ and R_f are expected asset return on asset i and the risk-free rate, respectively. The systematic risk measures, β_{im}, γ_{im} and δ_{im}, which are respectively the systematic covariance (*beta*), systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*), are defined as

$$\beta_{im} = \frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} = \frac{Cov(R_i, R_m)}{E[(R_m - E(R_m))^2]}, \tag{2.56}$$

$$\gamma_{im} = \frac{E[(R_i - E(R_i))(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} = \frac{Cos(R_i, R_m)}{E[(R_m - E(R_m))^3]}, \tag{2.57}$$

$$\delta_{im} = \frac{E[(R_i - E(R_i))(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]} = \frac{Cok(R_i, R_m)}{E[(R_m - E(R_m))^4]}. \tag{2.58}$$

Here, β_{im} is the covariance between the return on asset i and on the market portfolio ($Cov(R_i, R_m)$) divided by the variance of market portfolio return. γ_{im} is the *co-skewness* between the return on asset i and on the market portfolio ($Cos(R_i, R_m)$) divided by the third central moment of market portfolio return. δ_{im} is the *co-kurtosis* between the return on asset i and on the market portfolio ($Cok(R_i, R_m)$) divided by the fourth central moment of market portfolio return. In this thesis, we use the terminology of systematic risk measures: systematic covariance, systematic skewness and systematic kurtosis. Let c_1, c_2 and c_3, be the market prices or risk premiums for systematic covariance (β_{im}), systematic skewness (γ_{im}); systematic kurtosis (δ_{im}) in equation (2.55), respectively, which are given by

$$c_1 = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m), \tag{2.59}$$

$$c_2 = \frac{dE(W_F)}{dS(W_F)}S(R_m), \tag{2.60}$$

$$c_3 = \frac{dE(W_F)}{dK(W_F)}K(R_m). \tag{2.61}$$

Here, $\sigma(R_m)$, $S(R_m)$ and $K(R_m)$ are called standard deviation (*volatility*), skewness and kurtosis of the market portfolio return, R_m, respectively, which are defined as

$$\begin{aligned} \sigma(R_m) &= E[(R_m - E(R_m))^2]^{1/2}, \\ S(R_m) &= E[(R_m - E(R_m))^3]^{1/3}, \\ K(R_m) &= E[(R_m - E(R_m))^4]^{1/4}. \end{aligned} \tag{2.62}$$

Let $E(W_F)$, $\sigma(W_F)$, $S(W_F)$ and $K(W_F)$ be the expected value, standard deviation (*volatility*), skewness and kurtosis of an individual investor's end of period wealth, W_F, which are defined as

$$\begin{aligned} \sigma(W_F) &= E[(W_F - E(W_F))^2]^{1/2}, \\ S(W_F) &= E[(W_F - E(W_F))^3]^{1/3}, \\ K(W_F) &= E[(W_F - E(W_F))^4]^{1/4}. \end{aligned} \tag{2.63}$$

In equations (2.62)-(2.63), $\sigma^2(.)$, $S^3(.)$ and $K^4(.)$ are the second, third and fourth central moments of the market portfolio return, R_m and investor's end of period wealth, W_F, respectively. In the financial literature and throughout this thesis, $S(.)$ and $K(.)$ are called skewness and kurtosis, respectively, whereas in statistics, for example, skewness and kurtosis of the investor's end of period wealth, W_F are given by the scaled versions.

$$\begin{aligned}\sigma(W_F) &= E[(W_F - E(W_F))^2]^{1/2}, \\ S(W_F) &= E[(W_F - E(W_F))^3]/\sigma(W_F)^3, \\ K(W_F) &= E[(W_F - E(W_F))^4]/\sigma(W_F)^4.\end{aligned} \tag{2.64}$$

In the following section we derive the Four-Moment CAPM with utility function approximation, including the systematic risk measures, systematic covariance (β_{im}), systematic skewness (γ_{im}), and systematic kurtosis (δ_{im}) in equations (2.56)-(2.58), and the corresponding market prices, c_1, c_2 and c_3 in equations (2.59)-(2.61), respectively.

2.3.1.1 Derivation of Four-Moment CAPM using Utility Function

To derive the Four-Moment CAPM, we start to define the individual's expected utility function extending with skewness and kurtosis terms into that of quadratic form in equation (2.30) and then maximize it as a Lagrangian form. Next, we arrive from the individual equilibrium model to a market equilibrium model, and then the Four-Moment CAPM is obtained. Consider an individual investor's expected utility function[9] of the investor's wealth, $E[U(W_F)]$ which can be approximated by an fourth order Taylor series expansion, which is represented by

$$\begin{aligned}E[U(W_F)] &= U[E(W_F)] + \frac{1}{2!}U''[E(W_F)]\sigma(W_F)^2 \\ &+ \frac{1}{3!}U'''[E(W_F)]S(W_F)^3 + \frac{1}{4!}U''''[E(W_F)]K(W_F)^4.\end{aligned} \tag{2.65}$$

Note that under the assumption of normality, the expected utility is equivalent to a quadratic function (section 2.1), which is an exponential function of the mean and variance of wealth at the end of a period. More generally, without the assumption of normality, the utility function of investors is usually assumed to be the CARA (Constant Absolute Risk Aversion) utility (e.g. Levy and Markowitz (1979), Pulley (1981), Simaan (1993)).[10] For example, Jondeau and

[9] The individual investor's expected utility function of the investor's wealth, $E[U(W_F)]$ can be approximated by the n^{th} order Taylor series expansion defined in equation (2.29).

[10] Further details of special utility functions such as Constant Relative Risk Aversion (CRRA),

Rockinger (2004) showed that the fourth order Taylor expansion of the CARA utility function provided excellent performance for portfolio optimisation under the non-normality case. The CARA utility function is defined as follows.[11]

$$U(W_F) = -\exp(-\lambda W_F).$$

Here, λ denotes an absolute risk aversion. In the case of the CARA utility, the individual investor's expected utility function of the investor's wealth, $E(U(W_F))$, is represented by

$$E(U(W_F)) = -\exp(-\lambda E(W_F))\left[1 + \frac{\lambda^2}{2!}\sigma(W_F)^2 - \frac{\lambda^3}{3!}S(W_F)^3 + \frac{\lambda^4}{4!}K(W_F)^4\right].$$

To simplify the analysis, it is possible to show the link between the end of period wealth, W_F and the portfolio return, R_p, while assuming the initial wealth W_0 is 1 (equation (2.31)). The following equations are obtained.

$$E(W_F) = E(R_p) + 1, \tag{2.66}$$

$$\sigma(W_F) = \sigma(R_p), \tag{2.67}$$

$$S(W_F) = S(R_p), \tag{2.68}$$

$$K(W_F) = K(R_p). \tag{2.69}$$

Now to derive Four-Moment CAPM, we also need to display that

$$\sigma(W_F) = \sum_{i=1}^{N} x_i \beta_{ip} \sigma(R_p), \tag{2.70}$$

$$S(W_F) = \sum_{i=1}^{N} x_i \gamma_{ip} S(R_p), \tag{2.71}$$

$$K(W_F) = \sum_{i=1}^{N} x_i \delta_{ip} K(R_p), \tag{2.72}$$

and Decreasing Absolute Risk Aversion (DARA) can be found in Ziemann (2004) and Cuthbertson and Nitzsche (2005).

[11] This is primarily taken from Jondeau and Rockinger (2004).

where

$$\beta_{ip} = \frac{E[(R_i - E(R_i))(R_p - E(R_p))]}{E[(R_p - E(R_p))^2]}, \tag{2.73}$$

$$\gamma_{ip} = \frac{E[(R_i - E(R_i))(R_p - E(R_p))^2]}{E[(R_p - E(R_p))^3]}, \tag{2.74}$$

$$\delta_{ip} = \frac{E[(R_i - E(R_i))(R_p - E(R_p))^3]}{E[(R_p - E(R_p))^4]}, \tag{2.75}$$

are the systematic risk measures of asset i relative to portfolio. The proof requires us to show $\sum_{i=1}^{N} x_i\beta_{ip} = 1$, $\sum_{i=1}^{N} x_i\gamma_{ip} = 1$, and $\sum_{i=1}^{N} x_i\delta_{ip} = 1$. We have showed $\sum_{i=1}^{N} x_i\beta_{ip} = 1$ in equation (2.36). To display $\sum_{i=1}^{N} x_i\gamma_{ip} = 1$ and $\sum_{i=1}^{N} x_i\delta_{ip} = 1$, the same methods in equation (2.36) can be followed.

Now, the next step is to the maximization of individual investor's expected utility of the end of period wealth, W_F subject to a budget constraint, that is

$$\begin{aligned} \max_{\{x_0, x_1, \ldots, x_N\}} \quad & E[U(W_F)], \\ \text{subject to} \quad x_0 + \sum_{i=1}^{N} x_i &= 1. \end{aligned} \tag{2.76}$$

To maximize the individual investor's expected utility and solve the equilibrium condition in equation (2.76), a Lagrangian multiplier approach can be adopted. Define the function L as follows.

$$L = E[U(W_F)] - \lambda \left(x_0 + \sum_{i=1}^{N} x_i - 1 \right), \tag{2.77}$$

and note that from equation (2.65)

$$E[U(W_F)] = f(E(W_F),\, \sigma(W_F),\, S(W_F),\, K(W_F)),$$

as a function of $(E(W_F),\, \sigma(W_F),\, S(W_F),\, K(W_F))$. Hence,

$$L = f(E(W_F),\, \sigma(W_F),\, S(W_F),\, K(W_F)) - \lambda \left(x_0 + \sum_{i=1}^{N} x_i - 1 \right). \tag{2.78}$$

Here, λ is the Lagrangian multiplier and the expression in the brackets is equal to zero. Now we start by taking the first order total derivatives of the Lagrangian form in equation (2.78) with respect to proportions, x_0 and x_i, and setting these derivations equal to zero. Note that implicitly, these total derivatives, which use the fact (from equation (2.78)) that $E[U(W_F)]$ is a function of $E(W_F)$ as a function of $x_0, x_1, \ldots, x_N$ and $\sigma(W_F)$, $S(W_F)$ and $K(W_F)$ as functions of $x_1, \ldots, x_N$, are obtained from multivariate calculus as follows.

$$\begin{aligned}
\frac{\partial L}{\partial x_0} &= \frac{\partial E[U(W_F)]}{\partial E(W_F)}\frac{\partial E(W_F)}{\partial x_0} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\frac{\partial \sigma(W_F)}{\partial x_0} \\
&+ \frac{\partial E[U(W_F)]}{\partial S(W_F)}\frac{\partial S(W_F)}{\partial x_0} + \frac{\partial E[U(W_F)]}{\partial K(W_F)}\frac{\partial K(W_F)}{\partial x_0} - \lambda = 0, \\
&= \frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+R_f) - \lambda = 0,
\end{aligned} \tag{2.79}$$

using equations (2.41), (2.42) and

$$\frac{\partial S(W_F)}{\partial x_0} = 0, \quad \text{from equation (2.71)}, \tag{2.80}$$

$$\frac{\partial K(W_F)}{\partial x_0} = 0, \quad \text{from equation (2.72)}. \tag{2.81}$$

Similarly, the first derivatives of equation (2.78) with respect to x_i gives

$$\begin{aligned}
\frac{\partial L}{\partial x_i} &= \frac{\partial E[U(W_F)]}{\partial E(W_F)}\frac{\partial E(W_F)}{\partial x_i} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\frac{\partial \sigma(W_F)}{\partial x_i} \\
&+ \frac{\partial E[U(W_F)]}{\partial S(W_F)}\frac{\partial S(W_F)}{\partial x_i} + \frac{\partial E[U(W_F)]}{\partial K(W_F)}\frac{\partial K(W_F)}{\partial x_i} - \lambda = 0, \\
&= \frac{\partial E[U(W_F)]}{\partial E(W_F)}(1+E(R_i)) + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\beta_{ip}\sigma(R_p) \\
&+ \frac{\partial E[U(W_F)]}{\partial S(W_F)}\gamma_{ip}S(R_p) + \frac{\partial E[U(W_F)]}{\partial K(W_F)}\delta_{ip}K(R_p) - \lambda = 0,
\end{aligned} \tag{2.82}$$

using equations (2.44), (2.45) and

$$\frac{\partial S(W_F)}{\partial x_i} = \gamma_{ip}S(R_p), \quad \text{from equation (2.71)}, \tag{2.83}$$

$$\frac{\partial K(W_F)}{\partial x_i} = \delta_{ip}K(R_p), \quad \text{from equation (2.72)}. \tag{2.84}$$

Rearranging equation (2.79) and (2.82) gives

$$E(R_i) - R_f = -\frac{\frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}}{\frac{\partial E[U(W_F)]}{\partial E(W_F)}}\beta_{ip}\sigma(R_p) - \frac{\frac{\partial E[U(W_F)]}{\partial S(W_F)}}{\frac{\partial E[U(W_F)]}{\partial E(W_F)}}\gamma_{ip}S(R_p) - \frac{\frac{\partial E[U(W_F)]}{\partial K(W_F)}}{\frac{\partial E[U(W_F)]}{\partial E(W_F)}}\delta_{ip}K(R_p), \tag{2.85}$$

where $-\frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}/\frac{\partial E[U(W_F)]}{\partial E(W_F)}$, $-\frac{\partial E[U(W_F)]}{\partial S(W_F)}/\frac{\partial E[U(W_F)]}{\partial E(W_F)}$ and $-\frac{\partial E[U(W_F)]}{\partial K(W_F)}/\frac{\partial E[U(W_F)]}{\partial E(W_F)}$ equal to the investor's marginal rates of substitution (which is defined by minus the slope of the expected utility curve) between the expected and the standard deviation, skewness and kurtosis of end of period wealth, W_F, respectively. With same method in equations (2.47), (2.48), we now rearrange the investor's marginal rate of substitution. By knowing that through any point on the expected utility curve, $\frac{dE[U(W_F)]}{dE(W_F)} = 0$, because the expected utility, $E[U(W_F)] = f(E(W_F), \sigma(W_F), S(W_F), K(W_F))$ from equation (2.65), is constant while changing in expected return and variance of W_F are zero given the skewness and kurtosis of W_F, the following results are obtained.

$$\frac{dE[U(W_F)]}{dE(W_F)} = \frac{\partial E[U(W_F)]}{\partial E(W_F)} + \frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}\frac{d\sigma(W_F)}{dE(W_F)} = 0, \tag{2.86}$$

$$\frac{dE(W_F)}{d\sigma(W_F)} = -\frac{\frac{\partial E[U(W_F)]}{\partial \sigma(W_F)}}{\frac{\partial E[U(W_F)]}{\partial E(W_F)}}, \tag{2.87}$$

$$\frac{dE[U(W_F)]}{dE(W_F)} = \frac{\partial E[U(W_F)]}{\partial E(W_F)} + \frac{\partial E[U(W_F)]}{\partial S(W_F)}\frac{dS(W_F)}{dE(W_F)} = 0, \tag{2.88}$$

$$\frac{dE(W_F)}{dS(W_F)} = -\frac{\frac{\partial E[U(W_F)]}{\partial S(W_F)}}{\frac{\partial E[U(W_F)]}{\partial E(W_F)}}, \tag{2.89}$$

$$\frac{dE[U(W_F)]}{dE(W_F)} = \frac{\partial E[U(W_F)]}{\partial E(W_F)} + \frac{\partial E[U(W_F)]}{\partial K(W_F)}\frac{dK(W_F)}{dE(W_F)} = 0, \tag{2.90}$$

$$\frac{dE(W_F)}{dK(W_F)} = -\frac{\dfrac{\partial E[U(W_F)]}{\partial K(W_F)}}{\dfrac{\partial E[U(W_F)]}{\partial E(W_F)}}. \tag{2.91}$$

Substituting equations (2.87), (2.89) and (2.91) in equation (2.85) gives

$$\begin{aligned} E(R_i) - R_f &= \frac{dE(W_F)}{d\sigma(W_F)}\beta_{ip}\sigma(R_p) + \frac{dE(W_F)}{dS(W_F)}\gamma_{ip}S(R_p) \\ &+ \frac{dE(W_F)}{dK(W_F)}\delta_{ip}K(R_p). \end{aligned} \tag{2.92}$$

The final step is to define market case where the individual investor optimum portfolio is equivalent to the market portfolio; so, R_p is replaced by R_m. Then, equation (2.92) can be written as

$$\begin{aligned} E(R_i) - R_f &= \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m)\beta_{im} + \frac{dE(W_F)}{dS(W_F)}S(R_m)\gamma_{im} \\ &+ \frac{dE(W_F)}{dK(W_F)}K(R_m)\delta_{im}, \end{aligned} \tag{2.93}$$

where R_m, $E(R_i)$ and R_f are the market portfolio, expected asset i return and the risk-free rate, respectively. $\sigma(R_m)$, $S(R_m)$ and $K(R_m)$ are standard deviation, skewness and kurtosis of the market portfolio return, R_m, respectively. Here, β_{im}, γ_{im} and δ_{im} are the systematic risk measures of an asset i relative to market variance, skewness and kurtosis, such as systematic covariance (*beta*), systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*), respectively. Let c_1, c_2 and c_3 be given by

$$c_1 = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m), \quad c_2 = \frac{dE(W_F)}{dS(W_F)}S(R_m), \quad c_3 = \frac{dE(W_F)}{dK(W_F)}K(R_m),$$

which are the market prices for systematic risk measures in the Four-Moment CAPM defined earlier.

2.3.2 Three-Moment CAPM

When the individual investor's expected utility of the end of period wealth with a budget constraint (equation (2.65)) is independent of the kurtosis (i.e. $c_3 = 0$ in equation (2.55)); that is, the individual investor's expected utility function can be represented as the third order Taylor series expansion, the Four-Moment CAPM is reduced to the Three-Moment CAPM, developed by Kraus and Litzenberger (1976), identifies skewness term. It can be represented as

$$E(R_i) - R_f = c_1\beta_{im} + c_2\gamma_{im}, \tag{2.94}$$

where $E(R_i)$ and R_f are the expected asset return and the risk-free rate, respectively. Here, the systematic risk measures, systematic covariance, β_{im} and systematic skewness, γ_{im}, are defined as in equation (2.56) and (2.57), respectively. Let c_1 and c_2 be the market prices for β_{im} and γ_{im} from equations (2.59)-(2.60), defined as

$$c_1 = \frac{dE(W_F)}{d\sigma(W_F)}\sigma(R_m), \tag{2.95}$$

$$c_2 = \frac{dE(W_F)}{dS(W_F)}S(R_m). \tag{2.96}$$

2.3.3 Data Generating Processes for Higher-Moment CAPMs

To assess the necessity for the Higher-Moment CAPMs, Higher order Data Generating Processes (DGPs), namely the Quadratic Market Model and the Cubic Market Model, are discussed in the literature. For example, Barone-Adesi (1985) provides that the Quadratic Market Model is consistent with Kraus and Litzenberger's (1976) Three-Moment CAPM. Fang and Lai (1997) and Hwang and Satchell (1999) have also illustrated that the Higher order DGPs are consistent with their equivalent Higher-Moment CAPMs.

The Cubic Market Model is given by the following model

$$R_i - R_f = \kappa_i + \alpha_{1i}(R_m - R_f) + \alpha_{2i}(R_m - R_f)^2 + \alpha_{3i}(R_m - R_f)^3 + \varepsilon_i, \tag{2.97}$$

which is a simple polynomial regression of order three of the response, $R_i - R_f$ on the covariate $(R_m - R_f)$. The errors are assumed to be independent and identically distributed with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2$.

The Cubic Market Model (equation (2.97)) is consistent with Four-Moment CAPM (equation (2.55)). To show the link between the Four-Moment CAPM and Cubic Market Model, the systematic risk measures (β_{im}, γ_{im} and δ_{im}) can be expressed as

$$\begin{aligned}\beta_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} = \alpha_{1i} \\ &+ \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} \\ &+ \alpha_{3i}\frac{E[((R_m - R_f)^3 - E(R_m - R_f)^3)(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]},\end{aligned} \tag{2.98}$$

$$\begin{aligned}\gamma_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} = \alpha_{1i} \\ &+ \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} \\ &+ \alpha_{3i}\frac{E[((R_m - R_f)^3 - E(R_m - R_f)^3)(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]},\end{aligned} \tag{2.99}$$

$$\begin{aligned}\delta_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]} = \alpha_{1i} \\ &+ \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]} \\ &+ \alpha_{3i}\frac{E[((R_m - R_f)^3 - E(R_m - R_f)^3)(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]}.\end{aligned} \tag{2.100}$$

Note that the Four-Moment CAPM could only be employed if the DGP was at least cubic, that is, α_{3i} should be statistically significantly different from zero. If not, there will be collinearity in the systematic risk measures (β_{im}, γ_{im} and δ_{im}) of the Four-Moment CAPM.

Proof: For systematic covariance, β_{im}: Start by taking expectations of the Cubic Market Model, which gives

$$\begin{aligned} E(R_i - R_f) &= \kappa_i + \alpha_{1i} E(R_m - R_f) \\ &+ \alpha_{2i} E(R_m - R_f)^2 + \alpha_{3i} E(R_m - R_f)^3 + E(\varepsilon_i). \end{aligned} \tag{2.101}$$

Subtract from equation (2.97) to (2.101)

$$\begin{aligned} (R_i - R_f) - E(R_i - R_f) &= \alpha_{1i}((R_m - R_f) - E(R_m - R_f)) \\ &+ \alpha_{2i}\left((R_m - R_f)^2 - E(R_m - R_f)^2\right) \\ &+ \alpha_{3i}\left((R_m - R_f)^3 - E(R_m - R_f)^3\right) + \varepsilon_i, \\ R_i - E(R_i) &= \alpha_{1i}(R_m - E(R_m)) \\ &+ \alpha_{2i}\left((R_m - R_f)^2 - E(R_m - R_f)^2\right) \\ &+ \alpha_{3i}\left((R_m - R_f)^3 - E(R_m - R_f)^3\right) + \varepsilon_i. \end{aligned} \tag{2.102}$$

Multiply both sides of equation (2.102) by $R_m - E(R_m)$

$$\begin{aligned} (R_i - E(R_i))(R_m - E(R_m)) &= \alpha_{1i}(R_m - E(R_m))(R_m - E(R_m)) \\ &+ \alpha_{2i}\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m)) \\ &+ \alpha_{3i}\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m)) \\ &+ \varepsilon_i(R_m - E(R_m)). \end{aligned} \tag{2.103}$$

Take expected values both sides of equation (2.103)

$$\begin{aligned} &E\left[(R_i - E(R_i))(R_m - E(R_m))\right] \\ &= \alpha_{1i} E\left[(R_m - E(R_m))(R_m - E(R_m))\right] \\ &+ \alpha_{2i} E\left[\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))\right] \\ &+ \alpha_{3i} E\left[\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))\right]. \end{aligned} \tag{2.104}$$

Divide both sides in equation (2.104) by the variance of the market return, $E\left[(R_m - E(R_m))^2\right]$, which gives

$$\begin{aligned} \beta_{im} &= \frac{E\left[(R_i - E(R_i))(R_m - E(R_m))\right]}{E\left[(R_m - E(R_m))^2\right]} = \alpha_{1i} \\ &+ \alpha_{2i}\frac{E\left[\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))\right]}{E[(R_m - E(R_m))^2]} \\ &+ \alpha_{3i}\frac{E\left[\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))\right]}{E[(R_m - E(R_m))^2]}. \end{aligned} \tag{2.105}$$

For systematic skewness, γ_{im}: Similarly, multiply both side of equation (2.102) by $(R_m - E(R_m))^2$

$$\begin{aligned} &(R_i - E(R_i))(R_m - E(R_m))^2 \\ &= \alpha_{1i}(R_m - E(R_m))(R_m - E(R_m))^2 \\ &+\alpha_{2i}\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))^2 \\ &+\alpha_{3i}\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))^2 \\ &+\varepsilon_i(R_m - E(R_m))^2. \end{aligned} \tag{2.106}$$

Take expected values both sides of equation (2.106)

$$\begin{aligned} &E\left[(R_i - E(R_i))(R_m - E(R_m))^2\right] \\ &= \alpha_{1i}E\left[(R_m - E(R_m))(R_m - E(R_m))^2\right] \\ &+ \alpha_{2i}E\left[\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))^2\right] \\ &+ \alpha_{3i}E\left[\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))^2\right]. \end{aligned} \tag{2.107}$$

Divide both sides in equation (2.107) by the third central moment of the market return, $E\left[(R_m - E(R_m))^3\right]$, which gives

$$\begin{aligned} \gamma_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} = \alpha_{1i} \\ &+ \alpha_{2i}\frac{E[\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} \\ &+ \alpha_{3i}\frac{E[\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]}. \end{aligned} \tag{2.108}$$

For systematic kurtosis, δ_{im}: Finally, multiply both side of equation (2.102) by $(R_m - E(R_m))^3$

$$\begin{aligned} &(R_i - E(R_i))(R_m - E(R_m))^3 \\ = \quad & \alpha_{1i}(R_m - E(R_m))(R_m - E(R_m))^3 \\ + \quad & \alpha_{2i}\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))^3 \\ + \quad & \alpha_{3i}\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))^3 \\ + \quad & \varepsilon_i(R_m - E(R_m))^3. \end{aligned} \tag{2.109}$$

Take expected values both sides of equation (2.109)

$$\begin{aligned} &E[(R_i - E(R_i))(R_m - E(R_m))^3] \\ = \quad & \alpha_{1i}E[(R_m - E(R_m))(R_m - E(R_m))^3] \\ + \quad & \alpha_{2i}E[\left((R_m - R_f)^2 - E(R_m - R_f)^2\right)(R_m - E(R_m))^3] \\ + \quad & \alpha_{3i}E[\left((R_m - R_f)^3 - E(R_m - R_f)^3\right)(R_m - E(R_m))^3]. \end{aligned} \tag{2.110}$$

Divide both sides in equation (2.110) by the fourth central moment of the market return, $E\left[(R_m - E(R_m))^4\right]$, which gives

$$\begin{aligned} \delta_{im} \quad = \quad & \frac{E\left[(R_i - E(R_i))(R_m - E(R_m))^3\right]}{E[(R_m - E(R_m))^4]} = \alpha_{1i} \\ + \quad & \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]} \\ + \quad & \alpha_{3i}\frac{E[((R_m - R_f)^3 - E(R_m - R_f)^3)(R_m - E(R_m))^3]}{E[(R_m - E(R_m))^4]}. \end{aligned} \tag{2.111}$$

Provided that, the definition of systematic risk measures β_{im}, γ_{im} and δ_{im}, consistent with Four-Moment CAPM.

Throughout this thesis, the parameters of the Cubic Market Model must be estimated from more data, usually time series data of returns at regular intervals. Assuming that the parameters are constant over the whole time period being considered, then the excess return on asset i in period t $(R_{it} - R_{ft})$ might be assumed to be generated by the following model

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1i}(R_{mt} - R_{ft}) + \alpha_{2i}(R_{mt} - R_{ft})^2 + \alpha_{3i}(R_{mt} - R_{ft})^3 + \varepsilon_{it}, \tag{2.112}$$

where R_{it}, R_{ft} and R_{mt} are the returns on asset i, the risk-free rate and the market portfolio at time t ($t \in \{1, \ldots, T\}$). The errors are assumed to be independent and identically distributed with $\varepsilon_{it} \sim N(0, \sigma^2)$. Here, using $\hat{\kappa}_i$, $\hat{\alpha}_{1i}$, $\hat{\alpha}_{2i}$, and $\hat{\alpha}_{3i}$ from maximum likelihood estimation, and others, for example, $E\left((R_m - R_f)^2\right) \approx \frac{1}{T}\sum_{t=1}^{T}(R_{mt} - R_{ft})^2$; then, the systematic risk measures in equations (2.98), (2.99) and (2.100) can be estimated.

When $\alpha_{3i} = 0$ in equation (2.97), the Quadratic Market Model is given by the following model

$$R_i - R_f = \kappa_i + \alpha_{1i}(R_m - R_f) + \alpha_{2i}(R_m - R_f)^2 + \varepsilon_i, \tag{2.113}$$

which is a simple polynomial regression of order two of the response, $R_i - R_f$ on the covariate $(R_m - R_f)$. The errors are assumed to be independent and identically distributed with $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma^2$.

The Quadratic Market Model (2.113) is consistent with the Three-Moment CAPM (2.94). To show the link between the Three-Moment CAPM and Quadratic Market Model (see details in the Cubic Market Model), the systematic risk measures (β_{im} and γ_{im}) can be expressed as

$$\begin{aligned}
\beta_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]} = \alpha_{1i} \qquad (2.114)\\
&+ \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)\,(R_m - E(R_m))]}{E[(R_m - E(R_m))^2]},\\
\gamma_{im} &= \frac{E[(R_i - E(R_i))(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]} = \alpha_{1i} \qquad (2.115)\\
&+ \alpha_{2i}\frac{E[((R_m - R_f)^2 - E(R_m - R_f)^2)\,(R_m - E(R_m))^2]}{E[(R_m - E(R_m))^3]}.
\end{aligned}$$

Note that the Three-Moment CAPM could only be employed if the DGP was at least quadratic, that is, α_{2i} should be statistically significantly different from zero. If not, β_{im} and γ_{im} are equal.

Throughout this thesis, the parameters of the Quadratic Market Model must be estimated from more data, usually time series data of returns at regular intervals. Assuming that the parameters are constant over the whole time period being considered, then the excess return on asset i in period t $(R_{it} - R_{ft})$ might be assumed to be generated by the following model

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1i}(R_{mt} - R_{ft}) + \alpha_{2i}(R_{mt} - R_{ft})^2 + \varepsilon_{it}, \tag{2.116}$$

where R_{it}, R_{ft} and R_{mt} are the returns on asset i, the risk-free rate and the market portfolio at time t $(t \in \{1, \ldots, T\})$. The errors are assumed to be independent and identically distributed with $\varepsilon_{it} \sim N(0, \sigma^2)$. Here, using $\hat{\kappa}_i$, $\hat{\alpha}_{1i}$, and $\hat{\alpha}_{2i}$ from maximum likelihood estimation, and others, for example,
$E\left((R_m - R_f)^2\right) \approx \frac{1}{T}\sum_{t=1}^{T}(R_{mt} - R_{ft})^2$; then, the systematic risk measures in equations (2.114) and (2.115) can be estimated.

Chapter 3

Statistical Methodology

This chapter describes the statistical methodology used in this thesis, focusing on both time-invariant and time-varying coefficient models. In section 3.1, we review the linear model, and in section 3.2, we present a brief review of the additive model. In section 3.3, we describe a linear Gaussian state space model estimated via Kalman Filter approaches, and in section 3.4, we review the basic Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model and some of its extensions. In section 3.5, we discuss model selection and diagnostics. Throughout this thesis, vectors are denoted by bold type, while matrices are denoted by capital letters.

3.1 Linear Models

The basic form of the linear model is written as

$$Y_t = \boldsymbol{x}_t^{'}\boldsymbol{\beta} + \varepsilon_t, \qquad t = 1, \ldots, n, \tag{3.1}$$

where Y_t is a response variable which depends on covariates, $\boldsymbol{x}_t^{'}=(1, x_{t2}, \ldots, x_{tp})$. Here, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^{'}$ are unknown time-invariant coefficients, and $\{\varepsilon_t\}$ are independent and identically distributed random errors satisfying $\mathrm{E}(\varepsilon_t) = 0$ and $\mathrm{Var}(\varepsilon_t) = \sigma^2$ for $t = 1, \ldots, n$. In matrix notation, equation (3.1) is rewritten as

$$\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3.2}$$

where $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ is an $n \times 1$ random vector of responses; $X = (x_{tk})$ is an $n \times p$ design matrix of covariates; $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown time-invariant coefficients, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)$ is an $n \times 1$ random vector of errors with $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{Var}(\boldsymbol{\varepsilon}) = \sigma^2 I$. Hence, $\mathrm{E}(\boldsymbol{Y}) = X\boldsymbol{\beta}$ and $\mathrm{Var}(\boldsymbol{Y}) = \sigma^2 I$.

3.1.1 Estimation

A sensible approach is needed to estimate the value of $\boldsymbol{\beta}$ so that $X\boldsymbol{\beta}$ is as close as possible to $\boldsymbol{Y}$. One approach is to minimize the residual sum of squares (RSS) with respect to $\boldsymbol{\beta}$, which is known as the method of least squares. This method thus attempts to choose $\boldsymbol{\beta}$ by minimizing

$$Q = \sum_{t=1}^{n} \varepsilon_t^2 = \boldsymbol{\varepsilon}'\boldsymbol{\varepsilon} = (\boldsymbol{Y} - X\boldsymbol{\beta})'(\boldsymbol{Y} - X\boldsymbol{\beta}). \tag{3.3}$$

This can be achieved by expanding equation (3.3), which yields

$$Q = \boldsymbol{Y}'\boldsymbol{Y} - 2\boldsymbol{\beta}'X'\boldsymbol{Y} + \boldsymbol{\beta}'X'X\boldsymbol{\beta}. \tag{3.4}$$

Differentiating equation (3.4) with respect to $\boldsymbol{\beta}$ and setting it equal to zero gives

$$\frac{\partial Q}{\partial \boldsymbol{\beta}} = -2X'\boldsymbol{Y} + 2X'X\boldsymbol{\beta} = \mathbf{0}. \tag{3.5}$$

Dividing equation (3.5) by 2 and replacing $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}$, called the least squares estimator for $\boldsymbol{\beta}$, gives

$$X'X\hat{\boldsymbol{\beta}} = X'\boldsymbol{Y}. \tag{3.6}$$

This matrix form (equation (3.6)) is called the normal equations. Then, $\hat{\boldsymbol{\beta}}$ is obtained as

$$\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\boldsymbol{Y}, \tag{3.7}$$

as long as the design matrix X is of full rank, so that $(X^{'}X)^{-1}$ exists. The least squares estimator is unbiased i.e.

$$\begin{aligned} E(\hat{\boldsymbol{\beta}}) &= (X^{'}X)^{-1}X^{'}E(\boldsymbol{Y}) \\ &= (X^{'}X)^{-1}X^{'}X\boldsymbol{\beta} = I\boldsymbol{\beta} = \boldsymbol{\beta}, \end{aligned} \tag{3.8}$$

since $\mathrm{E}(\boldsymbol{Y}) = X\boldsymbol{\beta}$. The variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ is obtained as follows

$$\begin{aligned} \mathrm{Var}(\hat{\boldsymbol{\beta}}) &= (X^{'}X)^{-1}X^{'}\mathrm{Var}(\boldsymbol{Y})X(X^{'}X)^{-1} \\ &= \sigma^2(X^{'}X)^{-1}X^{'}X(X^{'}X)^{-1} = \sigma^2(X^{'}X)^{-1}, \end{aligned} \tag{3.9}$$

since $\mathrm{Var}(\boldsymbol{Y}) = \sigma^2 I$. The residual sum of squares is the smallest possible value of Q, obtained at $\hat{\boldsymbol{\beta}}$: $RSS = (\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})^{'}(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})$. An unbiased estimator for the variance σ^2 is

$$\hat{\sigma}^2 = \frac{RSS}{n-p} = \frac{(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})^{'}(\boldsymbol{Y} - X\hat{\boldsymbol{\beta}})}{n-p}, \tag{3.10}$$

where p is the number of parameters in the linear model.

3.1.2 Inference

To construct confidence intervals for $\boldsymbol{\beta}$ or to test a hypothesis related to the model, it is necessary to add an assumption about the distribution of the errors. Under the usual assumption of normality, the errors remain independent and identically distributed with mean zero and variance σ^2, so we assume that $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 I)$, which is equivalent to assuming $\boldsymbol{Y} \sim N(X\boldsymbol{\beta}, \sigma^2 I)$. Under these assumptions, it is straightforward to show that $\hat{\boldsymbol{\beta}}$ is also the maximum likelihood estimator for $\boldsymbol{\beta}$. It also follows that

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(X^{'}X)^{-1}). \tag{3.11}$$

To test a null hypothesis of the form $H_0 : \beta_i = C$ versus the alternative hypothesis, $H_A : \beta_i \neq C$ the test statistic is defined as follows.

$$T = \frac{\hat{\beta}_i - C}{\hat{\sigma}\,\sqrt{((X'X)^{-1})_{ii}}}. \tag{3.12}$$

Here, under the null hypothesis, T follows a t-distribution with $n-p$ degrees of freedom. Here, $\left((X'X)^{-1}\right)_{ii}$ is the i^{th} diagonal element of $(X'X)^{-1}$. Then, reject H_0 at the α level (e.g. $\alpha = 0.05$) if $|T| > t_{n-p,1-\alpha/2}$. Also, the *p-value* will be given by $Pr\left(|t_{n-p,1-\alpha/2}| > |T|\right)$. The $(1-\alpha)$ confidence interval for β_i is

$$\hat{\beta}_i \pm t_{n-p,1-\alpha/2}\,\hat{\sigma}\,\sqrt{((X'X)^{-1})_{ii}}. \tag{3.13}$$

For a comprehensive review of the linear model see Shumway and Stoffer (2006, Chap. 2), Faraway (2004, Chap. 2-3), Kutner et al. (2005, Chap. 5) and Wood (2006, Chap. 1).

3.2 Additive Models

The additive model, which was developed by Hastie and Tibshirani (1990), extends the linear model to include smooth functions of covariates whose shapes are estimated from the data rather than being specified by the investigator. The general form of an additive model with one additional covariate can be written as

$$Y_t = \boldsymbol{x}_t'\boldsymbol{\beta} + f(z_t) + \varepsilon_t, \quad t = 1, \ldots, n. \tag{3.14}$$

Here, the response Y_t is related to the covariates from the linear model $\boldsymbol{x}_t$ and a new covariate z_t . This new covariate has a potentially non-linear relationship with the response, which is represented by $f(.)$ and whose shape is estimated from the data. We note that equation (3.14) can be extended to have multiple covariates with non-linear shapes. However, as only the univariate case is used in this thesis, we do not discuss this possibility here. The function $f(z_t)$ can be constructed as a linear combination of $(m+1)^{th}$ order spline basis terms as follows.

$$f(z_t) = \sum_{i=1}^{r} B_i^m(z_t)\alpha_i. \tag{3.15}$$

Here, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$ are unknown parameters and $B_i^m(z_t)$, which is the i^{th} B-spline basis function of degree at m a point z_t, is the most popular choice for smoothing. Without considering the choice of knots, the B-spline basis functions are defined recursively as follows.

$$B_i^m(z_t) = \frac{z_t - z_{t\,i}}{z_{t\,i+m+1} - z_{t\,i}} B_i^{m-1}(z_t) + \frac{z_{t\,i+m+2} - z_t}{z_{t\,i+m+2} - z_{t\,i+1}} B_{i+1}^{m-1}(z_t), \quad i = 1, \ldots, r,$$

and

$$B_i^{-1}(z_t) = \begin{cases} 1 & \text{if } \; z_{t\,i} \leq z_t < z_{t\,i+1}, \\ 0 & \qquad \text{otherwise.} \end{cases}$$

Throughout this thesis, cubic spline ($m = 2$) is used (see Wood (2006)). For a comprehensive review of the properties of B-splines see de Boor (1978) and Eilers and Marx (1996).

To determine the degree of the smoothing for $f(z_t)$, a penalized regression spline approach can be adopted. This approach includes an overly large number of basis functions in equation (3.15) (*large* r), and minimizes excessive roughness via a penalty term. Estimation is achieved by minimizing the sum of squared residuals with a penalty term, that is minimizing

$$Q = (\boldsymbol{Y} - X\boldsymbol{\beta} - B\boldsymbol{\alpha})^{'} (\boldsymbol{Y} - X\boldsymbol{\beta} - B\boldsymbol{\alpha}) + \lambda \boldsymbol{\alpha}^{'} S \boldsymbol{\alpha}, \tag{3.16}$$

with respect to $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Here, B is a matrix of B-spline basis functions represented

$$B = \begin{bmatrix} B_1(z_1) & B_2(z_1) & \ldots & B_r(z_1) \\ B_1(z_2) & B_2(z_2) & \ldots & B_r(z_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_1(z_n) & B_2(z_n) & \ldots & B_r(z_n) \end{bmatrix}.$$

Here, we set $m = 2$ and drop the subscript m. $\boldsymbol{\alpha}$ is the vector of spline coefficients. S is a penalty matrix that is necessary to control the smoothness of $f(z_t)$ and reduce the effective degrees of freedom (EDF). Here, a common choice for penalty part is defined as

$$\sum_{i=2}^{r-1}(\alpha_{i-1} - 2\alpha_i + \alpha_{i+1})^2 = \boldsymbol{\alpha}' S \boldsymbol{\alpha}.$$

Here, the corresponding penalty matrix S is a second order random walk matrix given by

$$S = \begin{bmatrix} 1 & -2 & 1 & & & & \\ -2 & 5 & -4 & 1 & & & \\ 1 & -4 & 6 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 6 & -4 & 1 \\ & & & 1 & -4 & 5 & -2 \\ & & & & 1 & -2 & 1 \end{bmatrix}.$$

Let $X^* = (X, B)$ denote the combined matrix of X and B, the penalized least squares estimator of $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ is given by

$$\hat{\boldsymbol{\theta}} = (X^{*'} X^* + \lambda S^*)^{-1} X^{*'} \boldsymbol{Y}. \tag{3.17}$$

Here, $S^* = \begin{bmatrix} 0 & 0 \\ 0 & S \end{bmatrix}_{(p+r)\times(p+r)}$ and λ is a smoothing parameter, which determines the flexibility and smoothness of $\hat{f}(z_t)$. The data will be either under- or over-smoothed if λ is too low or too high. Many approaches exist to estimate λ, including the Akaike Information Criteria (AIC) (Hastie and Tibshirani (1990)), the Bayesian Information Criteria (BIC), the bias-corrected AIC ($AICc$) (Hurvich and Tsai (1989)) and the Generalized Cross Validation (GCV) (Wood (2006)). The aim is to minimize these criteria over all possible λ values and these are given by

$$\begin{aligned}
AIC(\lambda) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t(\lambda))^2 + 2tr(A(\lambda)), \qquad (3.18)\\
AICc(\lambda) &= \log\left(\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t(\lambda))^2\right) + 1 + \frac{2\,(tr(A(\lambda)) + 1)}{n - tr(A(\lambda)) - 2},\\
BIC(\lambda) &= \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t(\lambda))^2 + tr(A(\lambda))\log(n),\\
GCV(\lambda) &= \frac{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{\mu}_t(\lambda))^2}{(1 - {}^{tr(A(\lambda))}/_{n})^2}.
\end{aligned}$$

Here, $\hat{\mu}_t(\lambda)$ is the mean of Y_t and $A(\lambda)$ is the smoother matrix given by $A(\lambda) = X^*(X^{*'}X^* + \lambda S^*)^{-1}X^{*'}$, and $tr(A(\lambda))$ is defined as the effective degrees of freedom (EDF). For a comprehensive review of other types of the smoothing parameter λ selection criteria such as the UnBiased Risk Estimator ($UBRE$) (Craven and Wahba (1978)) and the REstricted Maximum Likelihood ($REML$) see Wood (2006). For a comprehensive review of penalized regression smoothers based on splines, cubic regression splines and penalized splines see Wood (2006).

3.3 Linear Gaussian State Space Models

3.3.1 Introduction

The linear state space form of a dynamic system with unobserved components is a powerful instrument for analyzing time series in many scientific areas such as engineering, physics, medicine and biostatistics (e.g. Hutchinson (1984), Roberts (1986), Zivot and Wang (2006) and Welch and Bishop (2006)). This general approach has also found numerous applications in the financial literature, with Harvey (1989), Harvey and Shephard (1993), Wells (1996), Durbin and Koopman (2001), Shumway and Stoffer (2006) and Mergner (2009) providing details about the use of the linear state space model for analyzing time series in finance.

The general form of the linear state space model with a normality assumption, which is referred to as a linear Gaussian state space model, consists of an observation or measurement equation and a state or transition equation. The observation equation can be written in the form

$$\boldsymbol{Y}_t = A_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \quad t = 1, \ldots, n. \tag{3.19}$$

Here, A_t is a $q \times p$ observation or measurement matrix that specifies how $\boldsymbol{\alpha}_t$, which is an unobserved $p \times 1$ state vector, can be converted into a $q \times 1$ vector of observations, $\boldsymbol{Y}_t$, at each time t. The observation equation is finished by assuming that the $q \times 1$ vector of errors, $\{\boldsymbol{\varepsilon}_t\}$, for $t = 1, \ldots, n$, are independent and identically distributed with $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, H_t)$, where H_t is a $q \times q$ variance matrix that is allowed to vary over time. The state equation can be modeled in the form

$$\boldsymbol{\alpha}_t = \Phi_t \boldsymbol{\alpha}_{t-1} + \boldsymbol{w}_t, \quad t = 1, \ldots, n, \tag{3.20}$$

which relates the state vector $\boldsymbol{\alpha}_t$ to its value $\boldsymbol{\alpha}_{t-1}$ at the previous time point. Here, the transition or speed parameter, Φ_t, is a $p \times p$ matrix, and $\{\boldsymbol{w}_t\}$ is assumed to be a $p \times 1$ vector of independent and identically distributed errors with $\boldsymbol{w}_t \sim N(\mathbf{0}, Q_t)$. Q_t is a $p \times p$ variance matrix of the errors in the state vector. The linear Gaussian state space model is completely specified by two further assumptions. Firstly, assuming that the initial state vector $\boldsymbol{\alpha}_0$, is Gaussian with mean, $\boldsymbol{\mu}_0$ and positive semi-definite variance matrix Σ_0. That is,

$$\boldsymbol{\alpha}_0 \sim N(\boldsymbol{\mu}_0, \Sigma_0), \tag{3.21}$$

and is independent of $\{\boldsymbol{w}_t\}$. Secondly, assuming that the observation and state errors $\{\boldsymbol{w}_t\}$ and $\{\boldsymbol{\varepsilon}_t\}$ are also independent of each other for all t.

The matrices, A_t, H_t, Φ_t and Q_t defined in equations (3.19) and (3.20) are referred to as system matrices. Typically, the system matrix A_t is assumed known, while (H_t, Q_t, Φ_t) are estimated from the data together with $(\boldsymbol{\mu}_0, \Sigma_0)$. For a comprehensive review of the linear Gaussian state space model see Harvey (1989, Chap. 3) and Durbin and Koopman (2001, Chap. 3). Throughout this thesis, the Kalman Filter, which was introduced by Kalman (1960) and Kalman and Bucy (1961), is used to estimate the parameters in the linear Gaussian state space model. In section 3.3.2 we present the Kalman Filter and Smoother algorithm used to estimate the unobserved state vector, $\boldsymbol{\alpha}_t$ and in section 3.3.3 we consider the estimation of the unknown hyperparameters using maximum likelihood. In section 3.3.4 the three forms of state space model used in this thesis are discussed.

3.3.2 The Kalman Filter and Smoother

The simplest Kalman Filter and Smoother algorithm is based on the following simplified model

$$\boldsymbol{Y}_t = A_t\boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, H), \tag{3.22}$$

$$\boldsymbol{\alpha}_t = \Phi\boldsymbol{\alpha}_{t-1} + \boldsymbol{w}_t, \qquad \boldsymbol{w}_t \sim N(\mathbf{0}, Q), \tag{3.23}$$

and initial state vector $\boldsymbol{\alpha}_0 \sim N(\boldsymbol{\mu}_0, \Sigma_0)$ for $t = 1, \ldots, n$. The elements of the system matrices $(H,\ Q,\ \Phi)$ are constant over time. Here, the primary aim of this analysis is to estimate the unobserved state vector, $\boldsymbol{\alpha}_t$, at time t given the available information Y_n=$\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n\}$ at time n. During this process, three types of problem are defined as follows: if $t > n$, this is a prediction problem, if $t = n$, this is a filtering problem, and if $t < n$, this is a smoothing problem. To achieve the solution to these problems, the Kalman Filter (*forward recursions*) and Smoother (*backward recursions*) are defined below. The Kalman Filter and Smoother algorithm is primarily taken from Shumway and Stoffer (2006, Chap. 6) where its proof can be found.

The conditional mean and variance of the state vector $\boldsymbol{\alpha}_t$ given data up to time n can be characterized as

$$\boldsymbol{\alpha}_t^n = E(\boldsymbol{\alpha}_t \mid Y_n), \tag{3.24}$$

$$P_t^n = \mathrm{Var}(\boldsymbol{\alpha}_t \mid Y_n). \tag{3.25}$$

To deal with the prediction $(t > n)$ and filtering $(t = n)$ problems, the forward recursion steps of the Kalman Filter and Smoother algorithm with initial conditions $\boldsymbol{\alpha}_0^0 = \boldsymbol{\mu}_0$ and P_0^0= Σ_0, can be implemented as follows, for $t = 1, \ldots, n$.

Prediction:

Set the state prediction

$$\boldsymbol{\alpha}_t^{t-1} = \Phi \boldsymbol{\alpha}_{t-1}^{t-1}. \tag{3.26}$$

Set the state variance prediction

$$P_t^{t-1} = \Phi P_{t-1}^{t-1} \Phi' + Q. \tag{3.27}$$

Filtering:

Set the innovations (*one-step ahead prediction error*)

$$\boldsymbol{v}_t = \mathbf{Y}_t - A_t \boldsymbol{\alpha}_t^{t-1}. \tag{3.28}$$

Set the variance matrices of the innovations

$$\Sigma_t = Var(\boldsymbol{v}_t) = A_t P_t^{t-1} A_t' + H. \tag{3.29}$$

Set the Kalman gain

$$K_t = P_t^{t-1} A_t' \Sigma_t^{-1}. \tag{3.30}$$

Set the state filtering

$$\boldsymbol{\alpha}_t^t = \boldsymbol{\alpha}_t^{t-1} + K_t \boldsymbol{v}_t. \tag{3.31}$$

Set the state variance filtering

$$P_t^t = P_t^{t-1} - K_t A_t P_t^{t-1}. \tag{3.32}$$

Cycle through equation (3.26) to (3.32) for each time t. Here, the forward recursion in equations (3.28) through (3.32) is called the Kalman Filter. For a comprehensive review of alternative derivations of the Kalman Filter see Harvey (1989, Chap. 3), Durbin and Koopman (2001, Chap. 4) and Mergner (2009, Chap. 3).

To deal with the smoothing ($t < n$) problem, the backward recursion of the Kalman Filter and Smoother is implemented, which starts with initial conditions $\boldsymbol{\alpha}_n^n$ (from equation (3.31)) and P_n^n (from equation (3.32)) obtained from the Kalman Filter with $t = n$. The Kalman Smoother works as follows. For $t = n, n-1, \ldots, 1$,

Set the smoothed state

$$\boldsymbol{\alpha}_{t-1}^n = \boldsymbol{\alpha}_{t-1}^{t-1} + J_{t-1}(\boldsymbol{\alpha}_t^n - \boldsymbol{\alpha}_t^{t-1}). \tag{3.33}$$

Set the smoothed error variance

$$P_{t-1}^n = P_{t-1}^{t-1} + J_{t-1}(P_t^n - P_t^{t-1})J_{t-1}', \tag{3.34}$$

where

$$J_{t-1} = P_{t-1}^{t-1}\Phi'[P_t^{t-1}]^{-1}. \tag{3.35}$$

This backwards recursion is referred to as a *"classical fixed interval smoother"*. It was introduced by Anderson and Moore (1979). For a comprehensive review of the derivation of the fast variants for state smoothing and disturbance smoothing, see Kohn and Ansley (1989), Jong (1991), Koopman (1993), Harvey (1989, Chap. 3), Durbin and Koopman (2001, Chap. 4), Fahrmeir and Tutz (2001, Chap. 8) and Mergner (2009, Chap. 3).

Throughout this thesis, to estimate a state $\boldsymbol{\alpha}_t$ given Y_n with $n > t$, we first apply the Kalman Filter (equations (3.28)-(3.32)) recursively until reaching the state $\boldsymbol{\alpha}_n$ and while moving forward we store the values $\boldsymbol{\alpha}_t^{t-1}$, $\boldsymbol{\alpha}_t^t$, P_t^{t-1} and P_t^t, $t = 1, \ldots, n$. Then, we move backwards by applying the Kalman Smoother (equations (3.33)-(3.35)) until reaching the state t, which we would like to estimate.

3.3.3 Estimation of Hyperparameters

Previously, we assumed that the system matrices H, Φ and Q and the initial mean $\boldsymbol{\mu}_0$ and variance Σ_0 are known, but we now consider the more usual situation in which at least some elements of the system matrices H, Φ and Q depend on a vector of unknown parameters, $\boldsymbol{\Theta}$, which are referred to as hyperparameters, that is,

$$\Phi = \Phi(\boldsymbol{\Theta}), \quad Q = Q(\boldsymbol{\Theta}), \quad H = H(\boldsymbol{\Theta}), \tag{3.36}$$

Here, we focus on the estimation of the vector of unknown parameters, $\boldsymbol{\Theta}$, by maximum likelihood. For the linear Gaussian state space model, we now derive the likelihood function which can be calculated by a routine application of the Kalman Filter based on the following assumptions (Durbin and Koopman (2001)): $\boldsymbol{\alpha}_0 \sim N(\boldsymbol{\mu}_0, \Sigma_0)$ where $\boldsymbol{\mu}_0$ and Σ_0 are known; $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n \sim N(\mathbf{0}, H)$; and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \sim N(\mathbf{0}, Q)$. For simplicity we continue to assume $\{\boldsymbol{w}_t\}$ and $\{\boldsymbol{\varepsilon}_t\}$ are uncorrelated. The likelihood function is

$$L_Y(\boldsymbol{\Theta}) = p(\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n; \boldsymbol{\Theta}) = p(\boldsymbol{Y}_1; \boldsymbol{\Theta}) \prod_{t=2}^{n} p(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta}), \tag{3.37}$$

where $p(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta})$ is the conditional density function of $\boldsymbol{Y}_t$ given the data set at time $t-1$; assuming that $\boldsymbol{\Theta}$ is the value of the unknown parameters vector. In practice the loglikelihood function is generally defined as

$$\log L_Y(\boldsymbol{\Theta}) = \sum_{t=1}^{n} \log p(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta}), \tag{3.38}$$

where $p(\boldsymbol{Y}_1 \mid \boldsymbol{Y}_0; \boldsymbol{\Theta}) = p(\boldsymbol{Y}_1; \boldsymbol{\Theta})$. For the linear Gaussian state space model in equations (3.22) and (3.23), it can be shown that the conditional distribution of $\boldsymbol{Y}_t$ is Gaussian with conditional mean

$$E(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta}) = A_t \boldsymbol{\alpha}_t^{t-1}, \tag{3.39}$$

and conditional variance matrix (equation (3.29))

$$\mathrm{Var}(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta}) = \Sigma_t. \tag{3.40}$$

Thus the conditional density function of $\boldsymbol{Y}_t$ is

$$p(\boldsymbol{Y}_t \mid \boldsymbol{Y}_{t-1}; \boldsymbol{\Theta}) = \frac{1}{2\pi^{q/2}} \mid \Sigma_t(\boldsymbol{\Theta}) \mid^{-1/2} \exp\left(-\frac{1}{2} \boldsymbol{v}_t(\boldsymbol{\Theta})' \Sigma_t(\boldsymbol{\Theta})^{-1} \boldsymbol{v}_t(\boldsymbol{\Theta})\right), \tag{3.41}$$

where $\boldsymbol{v}_t = \boldsymbol{Y}_t - A_t\boldsymbol{\alpha}_t^{t-1}$ (equation (3.28)). Substituting equation (3.41) into equation (3.38), the loglikelihood function can be written

$$\log L_Y(\boldsymbol{\Theta}) = -\frac{nq}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\log \mid \Sigma_t(\boldsymbol{\Theta}) \mid - \frac{1}{2}\sum_{t=1}^{n}\boldsymbol{v}_t(\boldsymbol{\Theta})'\Sigma_t(\boldsymbol{\Theta})^{-1}\boldsymbol{v}_t(\boldsymbol{\Theta}). \tag{3.42}$$

Here, $\boldsymbol{v}_t(\boldsymbol{\Theta})$ and $\Sigma_t(\boldsymbol{\Theta})$ are calculated routinely by the Kalman Filter (equation (3.26)-(3.32)). Also, we assume that $\Sigma_t(\boldsymbol{\Theta})$ is nonsingular for $t = 1, \ldots, n$. If this condition is not satisfied initially, the model can be redefined. This representation (equation (3.42)) of the loglikelihood function was first introduced by Schweppe (1965) and called prediction error decomposition by Harvey (1989, Chap. 3). For a comprehensive review of the loglikelihood function see Harvey (1989, Chap. 3), Durbin and Koopman (2001, Chap. 7) and Mergner (2009, Chap. 3).

The loglikelihood function (equation (3.42)) can be computed and maximized by numerical search algorithms to estimate the unknown parameter vector $\boldsymbol{\Theta}$. A Newton-Raphson method (also known as Newton's method) is the most widely used numerical search algorithm to update the unknown parameter vector $\boldsymbol{\Theta}$ by maximizing the loglikelihood function (equation (3.42)). The overall algorithm to update $\boldsymbol{\Theta}$ is primarily taken from Shumway and Stoffer (2006, Chap. 6) and defined as follows.

1. Start by selecting initial values (see details in Harvey (1989), Wells (1996)) for the unknown parameters vector, $\boldsymbol{\Theta}^{(0)}$.

2. Run the Kalman Filter using the initial values $\boldsymbol{\Theta}^{(0)}$ from step 1. Compute the set of innovations $\{\boldsymbol{v}_t^{(0)};\ t = 1, \ldots, n\}$ and the variance matrices of the innovations $\{\Sigma_t^{(0)};\ t = 1, \ldots, n\}$ used to calculate $\log L_Y(\boldsymbol{\Theta})$.

3. Run one iteration of the Newton-Raphson algorithm to update the estimates of $\boldsymbol{\Theta}$ to obtain a new set of estimates $\boldsymbol{\Theta}^{(1)}$.

4. Repeat steps 2 and 3 to obtain $\boldsymbol{\Theta}^{(j+1)}$ from $\boldsymbol{\Theta}^{(j)}$ and obtain the innovations $\{\boldsymbol{v}_t^{(j+1)};\ t = 1, \ldots, n\}$ and the variance matrices of the innovations $\{\Sigma_t^{(j+1)};\ t = 1, \ldots, n\}$ for $j = 1, 2 \ldots$. The production of the innovations also produces the estimates of the state vectors $\boldsymbol{\alpha}_1^{(j+1)}, \ldots, \boldsymbol{\alpha}_n^{(j+1)}$.

5. The algorithm stops when the value of $\boldsymbol{\Theta}^{(j+1)}$ differs from $\boldsymbol{\Theta}^{(j)}$, or when $L_Y(\boldsymbol{\Theta}^{(j+1)})$ differs from $L_Y(\boldsymbol{\Theta}^{(j)})$, by less than a predetermined small amount.

In the Newton-Raphson method, for a given value of $\boldsymbol{\Theta}$, the direction of the step in each iteration is determined by the gradient or score vector denoted as $\boldsymbol{g}(\boldsymbol{\Theta}) = {}^{\partial \log L_Y(\boldsymbol{\Theta})}/_{\partial \boldsymbol{\Theta}}$, while the size of the step is modified by the Hessian matrix denoted as $H(\boldsymbol{\Theta}) = {}^{\partial^2 \log L_Y(\boldsymbol{\Theta})}/_{\partial \boldsymbol{\Theta} \partial \boldsymbol{\Theta}'}$. Throughout this thesis the Newton-Raphson estimation of $\boldsymbol{\Theta}$ is accomplished using the *optim* package in the R software, using the BFGS (Broyden-Fletcher-Goldfarb-Shannon) method for solving an unconstrained nonlinear optimization problem numerically. Further details of the optimization with Newton's method and the BFGS method can be found in Fletcher (1987). For a comprehensive review of the gradient vector ($\boldsymbol{g}(\boldsymbol{\Theta})$) and the Hessian matrix ($H(\boldsymbol{\Theta})$) see Durbin and Koopman (2001, Chap. 7).

In the numerical optimization step, some elements of the unknown parameter vector $\boldsymbol{\Theta}$ are sometimes constrained. For example, the diagonal elements of the variance matrices Q and H are restricted to be positive, and the diagonal elements of the transition matrix Φ are restricted to the range $[0, 1]$. However, implementing these constraints within the numerical search algorithm (i.e. BFGS method) is inconvenient, and it is preferable that the maximization of the loglikelihood routines is performed with respect to unconstrained quantities. Therefore, the following transformations are defined while maximizing the loglikelihood function.

A diagonal element, σ^2, from the observation or state variance matrices H and Q, is restricted to be positive, so the loglikelihood function is maximized with respect to the unconstrained parameter

$$\Theta_\sigma = \log \sigma^2, \qquad -\infty < \Theta_\sigma < \infty. \tag{3.43}$$

A diagonal element, ϕ, of the transition matrix Φ is restricted to the range from zero to one, so the loglikelihood function is maximized with respect to the unconstrained parameter

$$\Theta_\phi = \sqrt{\frac{\phi}{1-\phi}}, \qquad -\infty < \Theta_\phi < \infty. \tag{3.44}$$

Further details of these transformations can be found in Wells (1996). For a comprehensive review of the initialization of the unknown parameter vector, $\boldsymbol{\Theta}$, while using these transformations within the maximization of the loglikelihood function via the Newton-Raphson method, see Harvey (1989), Wells (1996), Petris et al. (2009) and Mergner (2009).

3.3.4 Kalman Filter Based Models

Previously we presented the linear Gaussian state space model consisting of an observation equation (equation (3.22)) and a state equation (equation (3.23)). The observation equation can be seen as a time-varying coefficient regression model represented by

$$\boldsymbol{Y}_t = A_t \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0},\, H), \tag{3.45}$$

where the covariates A_t are multiplied by time-varying parameters $\boldsymbol{\alpha}_t$. For the state equation determining the evolution of $\boldsymbol{\alpha}_t$, three specific models are widely used in the financial literature. All of them are applied in this thesis and are outlined below. These specifications are defined and primarily taken from Wells (1996) and Mergner (2009).

Kalman Filter Random Coefficient

The Kalman Filter Random Coefficient (KFRC) model was first noted by Schaefer et al. (1975). The state equation of the time-varying coefficient regression model is written as

$$\boldsymbol{\alpha}_t = \bar{\boldsymbol{\alpha}} + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim N(\mathbf{0},\, Q), \tag{3.46}$$

where $\bar{\boldsymbol{\alpha}}$ is the mean of $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$. Here, $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$ are uncorrelated in time, and are globally smoothed towards a common mean $\bar{\boldsymbol{\alpha}}$. The set $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$ form a stationary sequence in time, with constant mean and variance.

Kalman Filter Random Walk

The Kalman Filter Random Walk (KFRW) model, which was first introduced by Samuelson (1965) in finance, has already been widely used in many scientific areas. The state equation of the time-varying coefficient model is written as

$$\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1} + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim N(\mathbf{0}, Q). \tag{3.47}$$

This first order random walk model assumes $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$ are autocorrelated, as $\boldsymbol{\alpha}_t$ equals $\boldsymbol{\alpha}_{t-1}$ plus random error. The set $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$ form a non-stationary sequence as Var $(\boldsymbol{\alpha}_t)$ increases with t.

Kalman Filter Mean Reverting

The Kalman Filter Mean Reverting (KFMR) model, which was proposed by Rosenberg (1973), is the primary model used in this thesis, because the first two models are special cases. The state equation of the time-varying coefficient regression model is written as

$$\boldsymbol{\alpha}_t - \bar{\boldsymbol{\alpha}} = \Phi(\boldsymbol{\alpha}_{t-1} - \bar{\boldsymbol{\alpha}}) + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \sim N(\mathbf{0}, Q). \tag{3.48}$$

To be a stationary sequence $\boldsymbol{\alpha}_0, \ldots, \boldsymbol{\alpha}_n$, the diagonal elements of the transition matrix, Φ should have modulus less than one. We note that these specifications are nested, because if $\Phi = \mathrm{I}$, KFMR becomes KFRW (equation (3.47)), while if $\Phi = 0$, KFMR becomes KFRC (equation (3.46)).

3.4 Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Models

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is widely used to model conditional (*time-varying*) volatility, which is a fundamental characteristic of financial time series. The GARCH-type models have also been applied in numerous empirical studies in finance, such as Engle (1982), Bollerslev (1986), and Glosten et al. (1993).

The rest of this section is outlined as follows. In section 3.4.1 we introduce the properties of GARCH-type models, and in section 3.4.2 we introduce the maximum likelihood estimation method in order to estimate the parameters of GARCH-type models.

3.4.1 Properties of GARCH-type Models

The basic form of the univariate time series model is written as

$$Y_t = \sigma_t z_t, \qquad t = 1, \ldots, n, \tag{3.49}$$

where $z_1, \ldots, z_t$ are independent and identically distributed with each $\mathrm{E}(z_t)$=0 and $\mathrm{Var}(z_t)$=1, and σ_t is the volatility that evolves over time. By definition, Y_t is serially uncorrelated with mean zero, but its conditional variance equals σ_t^2. There are several models proposed to specify the dynamic evolution of σ_t^2 in the literature. The following models, which have been widely used in the literature, are described and primarily taken from Peters (2001), Ruiz (2008), Mergner (2009) and Danielsson (2011).

ARCH Model

The Autoregressive Conditional Heteroskedasticity (ARCH) model, which is the simplest GARCH model, was introduced by Engle (1982). The conditional variance of the ARCH(p) model is given by

$$\sigma_t^2 = \mathrm{Var}(Y_t | Y_{t-1}, \ldots, Y_{t-p}) = \omega + \sum_{i=1}^{p} \psi_i Y_{t-i}^2, \quad t = p+1, \ldots, n, \tag{3.50}$$

in which p is the number of time lags over which the series is assumed to be autoregressive. The parameter restrictions for the ARCH(p) model are $\omega > 0$ and $\psi_i \geqslant 0$ for $i = 1, \ldots, p$, which are required to ensure positive conditional variance at every time t. Furthermore, $\sum_{i=1}^{p} \psi_i < 1$ is required to ensure stationarity of

σ_t^2. The conditional variance of the ARCH(1) model can be specified as

$$\sigma_t^2 = \omega + \psi_1 Y_{t-1}^2, \tag{3.51}$$

in which the current conditional variance, σ_t^2, depends on a constant and the response squared lagged by one time point. Suppose that the process is stationary and hence the unconditional variance of the ARCH(1) model, σ^2, is the same for all Y_t. Then,

$$E[Y_t] = E\left[E(Y_t|Y_{t-1})\right] = E[\sigma_t E(z_t|Y_{t-1})] = 0 \tag{3.52}$$

for all t. Therefore, as $E(Y_t|Y_{t-1}) = 0$

$$\begin{aligned} \sigma^2 = \text{Var}[Y_t] &= \text{Var}\left[E(Y_t|Y_{t-1})\right] + E\left[\text{Var}(Y_t|Y_{t-1})\right], \\ &= E\left[E(Y_t^2|Y_{t-1}) - E(Y_t|Y_{t-1})^2\right], \\ &= E\left[E(Y_t^2|Y_{t-1})\right] = E\left[\sigma_t^2 E(z_t^2|Y_{t-1})\right] = E\left[\sigma_t^2\right]. \end{aligned} \tag{3.53}$$

So, using (3.51),

$$\begin{aligned} \sigma^2 &= \omega + \psi_1 E[Y_{t-1}^2], \\ \sigma^2 &= \omega + \psi_1 \sigma^2, \\ \sigma^2 &= \frac{\omega}{1-\psi_1}. \end{aligned} \tag{3.54}$$

Here, for stationarity and finite unconditional variance σ^2, it is necessary for $\psi_1 < 1$.

GARCH Model

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of Bollerslev (1986) was created as an extension of the ARCH process to capture more complex volatility structure. The conditional variance of the GARCH(p, q) model is represented by

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \psi_i Y_{t-i}^2 + \sum_{j=1}^{q} \theta_j \sigma_{t-j}^2, \quad t = min(p, q) + 1, \ldots, n, \tag{3.55}$$

where parameter restrictions, $\omega > 0$, $\psi_i \geqslant 0$, and $\theta_i \geqslant 0$ are required to ensure positive conditional variance at every t, these restrictions have been derived by Nelson and Cao (1992). Furthermore, $\sum_{i=1}^{p} \psi_i + \sum_{j=1}^{q} \theta_i < 1$ is required to ensure covariance stationarity. The conditional variance of the GARCH(1,1) model is defined as

$$\sigma_t^2 = \omega + \psi_1 Y_{t-1}^2 + \theta_1 \sigma_{t-1}^2, \tag{3.56}$$

in which current conditional variance, σ_t^2, depends on a constant, and both the response and conditional variance lagged by one time point. Under the assumption of covariance stationarity, the unconditional variance of GARCH(1,1) model can be represented as

$$\begin{aligned} \sigma^2 \equiv E[\sigma_t^2] &= \omega + \psi_1 E[Y_{t-1}^2] + \theta_1 E[\sigma_{t-1}^2], \\ \sigma^2 &= \omega + \psi_1 \sigma^2 + \theta_1 \sigma^2, \\ \sigma^2 &= \frac{\omega}{1 - \psi_1 - \theta_1}, \end{aligned} \tag{3.57}$$

Therefore, for stationarity and finite unconditional variance σ^2, it is required that $\psi_1 + \theta_1 < 1$.

GJR-GARCH Model

The GJR-GARCH(p, q) model, which was developed by Glosten et al. (1993), is also widely used to capture the leverage effect with an indicator variable. The leverage effect, which was first described by Black (1976), refers the asymmetric response of the volatility to positive and negative movements in financial time series. The conditional variance of the GJR-GARCH(p, q) model is defined as

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} (\psi_i Y_{t-i}^2 - \zeta_i I_{t-i} Y_{t-i}^2) + \sum_{j=1}^{q} \theta_j \sigma_{t-j}^2, \tag{3.58}$$

where ζ_i shows the leverage term. Also, I_{t-i} is an indicator variable that takes the value 1 if the t^{th} return is negative or zero and the value 0 otherwise. Let the conditional variance of the GJR-GARCH(1,1) model be represented as

$$\sigma_t^2 = \omega + \psi_1 Y_{t-1}^2 + \zeta_1 I_{t-1} Y_{t-1}^2 + \theta_1 \sigma_{t-1}^2, \tag{3.59}$$

where I_{t-1} takes the value 1 for $Y_{t-1} \leqslant 0$ and 0 otherwise. Note that for symmetrically distributed z_t (3.49), the process is covariance stationary if $\psi_1 + \theta_1 + \frac{1}{2}\zeta_1 < 1$.

3.4.2 Maximum Likelihood Estimation

To estimate the unknown parameters of GARCH-type models maximum likelihood estimation (MLE) method can be employed. Throughout this thesis two error distributions, the normal and the student-t, are considered in the MLE which are taken from Peters (2001), Christoffersen (2003) and Danielsson (2011).

Normal Distribution

The normal distribution is the one that is most widely used to estimate and forecast GARCH-type models. In this case, the observed return at time t is defined as

$$Y_t = \sigma_t z_t \quad \text{with} \quad z_t \sim N(0,1), \quad t = 1, \ldots, n. \tag{3.60}$$

Note that the conditional variance σ_t starts at $t = 2$, since we do not know Y_0. The joint likelihood function is thus

$$L = \prod_{t=2}^{n} l_t = \prod_{t=2}^{n} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{Y_t^2}{2\sigma_t^2}\right), \tag{3.61}$$

but it is the loglikelihood function

$$\begin{aligned} \log L &= \sum_{t=2}^{n} \log(l_t), \\ &= \sum_{t=2}^{n} \left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{1}{2}\frac{Y_t^2}{\sigma_t^2}\right]. \end{aligned} \tag{3.62}$$

that is maximised instead.

Student-t Distribution

The student-t distribution is generally used for capturing heavy tails when estimating and forecasting GARCH-type models. In this case, the observed return at time t is defined as

$$Y_t = \sigma_t z_t \quad \text{with} \quad z_t \sim t_v, \quad t = 1, \ldots, n. \tag{3.63}$$

The same issue as before arises with unknown Y_0. Therefore, the density function starts at $t = 2$, and the joint likelihood function is

$$L = \prod_{t=2}^{n} l_t = \prod_{t=2}^{n} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)\sigma_t^2}} \left(1 + \frac{1}{(v-2)}\frac{Y_t^2}{\sigma_t^2}\right)^{-\frac{v+1}{2}}. \tag{3.64}$$

As before the logarithm of the joint likelihood function is maximised. Here, v denotes the degrees of freedom and satisfies $v > 2$, and $\Gamma(.)$ is the gamma function. The lower value of v, the fatter the tails, and as $v \longrightarrow \infty$, this distribution tends to the standard normal.

For the GARCH(1,1) model (3.56) with a normal likelihood function the parameter estimation algorithm is as follows (Danielsson (2011)). Other models utilise similar algorithms, and are not described here.

1. Initialise the algorithm by setting

$$\sigma_1^2 = \frac{1}{n}\sum_{t=1}^{n} Y_t^2, \tag{3.65}$$

as $E(Y_t) = 0$.

2. Create a grid of pairs $\left(\psi_1^{(j)}, \theta_1^{(j)}\right)$ of candidate values satisfying

   (a) $\psi_1^{(j)} + \theta_1^{(j)} < 1$,

   (b) $\psi_1^{(j)}, \theta_1^{(j)} \geq 0$.

3. For each candidate pair $\left(\psi_1^{(j)}, \theta_1^{(j)}\right)$ compute $\omega^{(j)}$ from (3.57) (which re-

duces the number of parameters estimated in the model by one) assuming that $\sigma_1^2 = \sigma^2$. Then using (3.56) compute $\sigma_2^2, \dots, \sigma_n^2$. Then evaluate the loglikelihood function (3.62).

4. Choose $\left(\psi_1^{(j)}, \theta_1^{(j)}\right)$ that maximize (3.62) overall pairs of values considered as the MLE.

3.5 Model Selection and Diagnostics

Previously Linear (3.1), Additive (3.2), State Space (3.3) and GARCH (3.4) models have been discussed, but after a model has been implemented, it is necessary to check whether the assumptions underlying the model hold. By applying various statistical tests and graphical procedures, we can evaluate the quality of the fitted model, and a sample of these techniques are described in sections 3.5.1 and 3.5.2.

3.5.1 Model Selection

While many criteria for comparing the fit of multiple models have been developed, we only focus on six: R^2, *Adjusted* R^2, the Mean Absoulate Error (MAE), the Mean Square Error (MSE), the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Further details of these and other criteria are given in Faraway (2004) and Kutner et al. (2005).

The coefficient of determination or the percentage of variance explained, R^2, is one of the most widely used measures of how well the model fits the data. It is defined as

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2} = 1 - \frac{RSS}{TSS}, \tag{3.66}$$

where $\bar{Y}$ denotes the mean of the data and $\hat{Y}_t$ is the fitted value for the t^{th} ($t = 1, \dots, n$) unit. Here, RSS and TSS are the residual sum of squares and the total sum of squares, respectively. Note that adding a predictor to a model can only decrease the RSS and so only increase the R^2. Hence, R^2 by itself is not a good criterion, because it always chooses the largest possible model. To penalize models having large number of predictors, another commonly used model selection criterion is *Adjusted* R^2 and is denoted as $\bar{R}^2$. It is represented by

$$\bar{R}^2 = 1 - \left(\frac{n-1}{n-p}\right)\frac{RSS}{TSS}, \tag{3.67}$$

where n is the number of data points and p is the number of estimated predictors. Models with larger *Adjusted* R^2 values indicates better fit.

The fit of models is compared using two different summaries of the errors, the Mean Absolute Error (MAE) and the Mean Squared Error (MSE). The MAE is defined as

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{Y}_t - Y_t|. \tag{3.68}$$

Here, a potential problem with using MAE is it weighs all errors equally. Alternative approach is the MSE represented as

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\hat{Y}_t - Y_t)^2. \tag{3.69}$$

Here, the use of a squared term in the equation places a heavier penalty on outliers than the MAE. Using MAE and MSE in out-of-sample procedures gives a measure of the forecasting ability of the models. According to these two measures of forecasting error, the models with lowest MAE and MSE values indicate better forecasting performance.

Another way of comparing different models is to compare the loglikelihood from the fitted model, denoted by $\log L_Y$, with the corresponding loglikelihood values from competing models. In general, the larger number of predictors that a model contains the larger its loglikelihood. To penalize models having large number of predictors, two popular alternatives, the Akaike information criterion (*AIC*, Akaike (1974)) and the Bayesian information criterion (*BIC*, Schwarz (1978)) have been proposed. They are defined in terms of the loglikelihood function as follows.

$$AIC = -2\log L_Y + 2w, \tag{3.70}$$

$$BIC = -2\log L_Y + w\log n, \tag{3.71}$$

where w is the number of predictors in the regression model, and L_Y is the maximized value of the likelihood function for the fitted model. Note that in the special case of least squares estimation with normally distributed errors, $-2\log L_Y \approx n(1+\log(2\pi))+n\log(\hat{\sigma}^2)$. Here, $\hat{\sigma}^2 = {}^{RSS}/_n$ is the estimated variance of the residuals after fitting the model. In model comparisons on the same data, the constant can be ignored, because it is the same for a given data set and assumed error distribution (Faraway (2004)). In theory, the smaller AIC and BIC values, the better fitting the model is to the data.

3.5.2 Model Diagnostics

3.5.2.1 Univariate Model Diagnostics

Univariate diagnostic procedures are intended to check how well the assumptions of the regression model are satisfied. These assumptions underlying the regression model are that the residuals are normally distributed, independent and have constant variance. Throughout this thesis, these assumptions need to be checked using various univariate diagnostic tests and graphical procedures based on the residuals. These techniques are primarily taken from Harvey (1989), Durbin and Koopman (2001) and Faraway (2004). The simple raw residuals can be represented as

$$r_t = Y_t - \hat{Y}_t, \tag{3.72}$$

where Y_t is the observed response and $\hat{Y}_t$ is the fitted value for the t^{th} $(t = 1, \ldots, n)$ unit. The problem of using the residuals $\{r_t\}_{t=1}^{n}$ is that their variances may differ, so detecting outliers is difficult. The problem is overcome using standardised residuals defined as

$$s_t = \frac{Y_t - \hat{Y}_t}{\sqrt{\mathrm{Var}(Y_t - \hat{Y}_t)}}. \tag{3.73}$$

The standardised residuals $\{s_t\}_{t=1}^n$ will have approximately $N(0,1)$ distributions if the linear Gaussian model holds. Data points with standardised residuals which are unusually large relative to $N(0,1)$ (e.g. $\mid s_t \mid > 3$ or $\mid s_t \mid > 4$) may be considered potential outliers in the sense that these Y_t's are much farther away from their $\hat{Y}_t$'s.

Normality

To test normality of the residuals, the Jarque-Bera test (*JB*, Jarque and Bera (1980)) is a goodness-of-fit test of whether the skewness and kurtosis of the data are appropriate for a Gaussian distribution. The sample skewness and kurtosis of the standardised residuals are given by

$$S = \frac{\frac{1}{n}\sum_{t=1}^{n}(s_t - \bar{s})^3}{\left(\frac{1}{n}\sum_{t=1}^{n}(s_t - \bar{s})^2\right)^{3/2}}, \quad K = \frac{\frac{1}{n}\sum_{t=1}^{n}(s_t - \bar{s})^4}{\left(\frac{1}{n}\sum_{t=1}^{n}(s_t - \bar{s})^2\right)^{2}}, \tag{3.74}$$

where $\bar{s}$ is the mean of the standardised residuals, $\{s_t\}_{t=1}^n$. The Jarque-Bera (JB) test statistic is defined as

$$JB = n\left\{\frac{S^2}{6} + \frac{(K-3)^2}{24}\right\}. \tag{3.75}$$

Here, under the null hypothesis of normality, JB follows a chi-squared distribution with 2 degrees of freedom. Alternatively, the standardised residuals can be assessed for normality using a *Q-Q* plot. This plot compares the standardised residuals to ideal normal observations. The ordered standardised residuals are plotted against $\Phi^{-1}\left(\frac{t}{n+1}\right)$ for $t = 1, \ldots, n$. In a *Q-Q* plot a straight line suggests the distributional assumption is appropriate, while a curved line suggests that a heavier or lighter tailed distribution is appropriate.

Heteroskedasticity

To test for heteroskedasticity (non-constant variance), the simplest diagnostic test statistic is defined (Durbin and Koopman (2001)) by

$$Het(h) = \frac{\sum\limits_{t=n-h+1}^{n} s_t^2}{\sum\limits_{t=1}^{h} s_t^2}. \tag{3.76}$$

Here, given the null hypothesis of homoscedasticity (constant variance), $Het(h)$ follows a $F_{h,h}$ distribution for some preset positive integer h which is the nearest integer to $n/3$. In contrast, a graphical technique is to plot the standardised residuals against the fitted values and non-constant variance is apparent by any striking pattern.

Time Series Autocorrelation

To test for temporal autocorrelation, the Ljung-Box test is widely used, which was developed by Ljung and Box (1978), and is often called a portmanteau test. The test statistic is

$$LB(L) = n(n+2)\sum_{k=1}^{L} \frac{\rho_k^2}{n-k}, \tag{3.77}$$

where L is the number of lags being tested and ρ_k is the sample autocorrelation of the standardised residuals at lag k defined as

$$\rho_k = \frac{\sum\limits_{t=k+1}^{n} (s_t - \bar{s})(s_{t-k} - \bar{s})}{\sum\limits_{t=1}^{n} (s_t - \bar{s})^2}, \qquad k = 1, 2, \ldots . \tag{3.78}$$

Here, given the null hypothesis of no autocorrelation, $LB(L)$ follows a chi-squared distribution with L degrees of freedom. Also, the plot of the sample autocorrelation (3.78) of $\{s_t\}_{t=1}^{n}$ can be visually assessed to check for the presence of autocorrelation.

3.5.2.2 Multivariate Model Diagnostics

Multivariate diagnostic procedures are intended to check how well the assumptions of the multivariate regression model are satisfied. The assumptions underlying the multivariate regression model are that the residuals are multivariate normally distributed, multivariate independent and have multivariate constant variance. These assumptions need to be checked using various multivariate diagnostic tests and graphical procedures based on the residuals in this thesis. Note that the grapical procedures which are discussed in section 3.5.2.1 for checking the regression model assumptions in the univariate context can be extended to the multivariate context. Here, multivariate diagnostic tests are only represented and primarily taken from Harvey (1989), Pfaff (2008) and Mahdi (2011). The raw residuals vector can be defined as

$$\boldsymbol{r}_t = \boldsymbol{Y}_t - \hat{\boldsymbol{Y}}_t, \tag{3.79}$$

Here, $\boldsymbol{Y}_t$ is a $q \times 1$ observed response vector and $\hat{\boldsymbol{Y}}_t$ is a $q \times 1$ fitted value vector for the t^{th} $(t = 1, \ldots, n)$ unit. The standardized residuals vector is defined as

$$\boldsymbol{s}_t = P_t^{-1}\boldsymbol{r}_t. \tag{3.80}$$

Here, P_t is a $q \times q$ lower triangular matrix for the t^{th} $(t = 1, \ldots, n)$ unit (its diagonal elements are positive) such that $P_t P_t^{'} = \Sigma_{\boldsymbol{r}_t}$. Here, $\Sigma_{\boldsymbol{r}_t}$ is the residual variance matrix for the t^{th} unit (Pfaff (2008)).

Multivariate Normality

To test for multivariate normality, the multivarite Jarque-Bera test is a goodness-of-fit test of whether the multivariate skewness and multivariate kurtosis of the data are appropriate for a multivariate Gaussian distribution. Here, the centered standardized residuals vector is used and defined as

$$\boldsymbol{cs}_t = P_t^{-1}(\boldsymbol{r}_t - \bar{\boldsymbol{r}}_t). \tag{3.81}$$

The multivariate skewness (S) is defined as

$$\begin{aligned} S_i &= \frac{\frac{1}{n}\sum_{t=1}^{n}(cs_{it})^3}{\left(\frac{1}{n}\sum_{t=1}^{n}(cs_{it})^2\right)^{3/2}}, \qquad i = 1, \dots, q, \\ S &= (S_1, \dots, S_q)(S_1, \dots, S_q)'. \end{aligned} \tag{3.82}$$

The multivariate kurtosis (K) is defined as

$$\begin{aligned} K_i &= \frac{\frac{1}{n}\sum_{t=1}^{n}(cs_{it})^4}{\left(\frac{1}{n}\sum_{t=1}^{n}(cs_{it})^2\right)^{2}}, \qquad i = 1, \dots, q, \\ K &= (K_1 - 3, \dots, K_q - 3)(K_1 - 3, \dots, K_q - 3)'. \end{aligned} \tag{3.83}$$

The multivariate Jarque-Bera (MJB) test statistic is defined as

$$MJB = n\left\{\frac{S}{6} + \frac{K}{24}\right\}. \tag{3.84}$$

Here, under the null hypothesis of multivariate normality, MJB follows a chi-squared distribution with $2q$ degrees of freedom. In addition, the multivariate skewness and the multivariate kurtosis tests follow a chi-squared distribution with q degrees of freedom (Pfaff (2008)).

Multivariate Heteroskedasticity

To test for multivariate heteroskedasticity, the simplest diagnostic test statistic defined in 3.76 which can be extended to the multivariate context as follows.

$$MHet(hq) = \frac{\sum_{j=(nq)-(hq)+1}^{nq} v_j^2}{\sum_{j=1}^{hq} v_j^2}. \tag{3.85}$$

Here, $\boldsymbol{v}$ is a $nq \times 1$ vector of the ordered standardized residuals s_{it} ($t = 1, \dots, n$; $i = 1, \dots, q$). Here, given the null hypothesis of multivariate homoscedasticity (constant variance), $MHet(hq)$ follows a $F_{hq,hq}$ distribution for some preset positive integer h, which itself is the nearest integer to $n/3$.

Multivariate Time Series Autocorrelation

To test for multivariate temporal autocorrelation, the multivariate extension of the Ljung-Box test (multivariate portmanteau test) is proposed by Hosking (1980). The test statistic for the multivariate portmanteau test is defined as

$$MLB(L) = n \sum_{k=1}^{L} tr\left(C_k^{\prime} C_0^{-1} C_k C_0^{-1}\right). \tag{3.86}$$

Here, $C_k = \dfrac{1}{n} \sum_{t=k+1}^{n} \boldsymbol{s}_t \boldsymbol{s}_{t-k}^{\prime}$ and L is the number of lags being tested. Here, given the null hypothesis of no multivariate autocorrelation, $MLB(L)$ follows a chi-squared distribution with $q^2L - m$ degrees of freedom where m is the total number of estimated parameters (Harvey (1989)).

Chapter 4

Modelling the Time-varying Systematic Covariance Risk of Turkish Industry Sector Portfolios

4.1 Introduction

Financial researchers have utilised systematic risk measures to provide guidelines for testing asset pricing models, determining risk, and financial investment decisions such as portfolio choice and capital budgeting over recent decades. The most widely used systematic risk measure, systematic covariance risk, is summarised by the parameter *beta* in the Two-Moment Capital Asset Pricing Model (CAPM) of Sharpe-Lintner-Mossin (1960s). In the context of this CAPM, the systematic covariance (*beta*) risk is stable over time and is commonly estimated via Ordinary Least Squares. However, there is now substantial empirical evidence in the recent literature (e.g. Brooks et al. (1998), Faff et al. (2000), Mergner and Bulla (2008) and Choudhry and Wu (2009)) that the Two-Moment CAPM with constant systematic covariance risk may be misleading and insufficient for modelling and forecasting the returns of financial data. The inadequacies of the Two-Moment CAPM have encouraged financial researchers to explore the stochastic behaviour of the systematic covariance risk in the Two-Moment CAPM.

One extension in this vein is to allow the systematic covariance (*beta*) risk to change linearly over time in the Two-Moment CAPM, which is often known as the conditional Two-Moment CAPM. Recently, several papers have extensively investigated the instability of the systematic covariance risk for different countries and firms, by comparing the modelling and forecasting abilities of the unconditional and conditional Two-Moment CAPMs. Examples include financial data relating to Australia (Faff et al. (1992), Brooks et al. (1992), Brooks et al. (1998), Brooks et al. (2002)), the United Kingdom (Faff et al. (2000)), firms within the UK (Choudhry and Wu (2009)), Europe (Wells (1994), Mergner and Bulla (2008)), the USA (Fabozzi and Francis (1978), Sunder (1980), Bos and Newbold (1984), Kim (1993)), and Turkey (Odabasi (2003)). Most of this research has modelled developed countries and firms systematic covariance risks, and only a few papers have focused on emerging countries and firms. Because of this lack of emerging market research, we focus this chapter on confirming the evidence for the instability of the systematic covariance risk in Turkish industry sector portfolios, which are classified among the emerging markets (`http://www.msci.com`). The most well-known modelling techniques for assessing this instability are GARCH-type models and Kalman Filter based approaches, which have been applied in many studies with differing results. In this chapter we apply these techniques to Turkish industry sector portfolios data, and undertake the comparison of the modelling and forecasting performance of these modelling techniques.

The first technique for modelling and forecasting time-varying systematic covariance risk are GARCH-type models (Engle (1982) and Bollerslev (1986)), which are based on estimating the conditional variance and covariance between asset and market portfolio returns. In this chapter the most widely used GARCH-type models, the standard GARCH (Bollerslev (1986)) and Glosten-Jagannathan-Runkle GARCH (GJR-GARCH, a non-linear extension of the GARCH model, (Glosten et al. (1993)) have been applied. GJR-GARCH models can capture asymmetric effects of the conditional volatility of negative and positive shocks on returns. In the literature the performance of these models has been compared using in-sample and out-of-sample procedures. For example, Mergner and Bulla (2008) compared GARCH-type models, such as GARCH and GJR-GARCH, for the time-varying behaviour of systematic covariance risk for 18 pan-

European sectors using weekly data over the period December 1987 to February 2005, and assessed both in-sample modelling and out-of-sample forecasting performance. Also, Choudhry and Wu (2009) estimated the weekly time-varying systematic covariance risk of UK firms from January 1989 to December 2003 using a GJR-GARCH model and a bivariate GARCH, Baba-Engle-Kraft-Kroner GARCH (BEKK-GARCH, Engle and Kroner (1995)), and forecasts of the time-varying betas were examined to evaluate out-of-sample forecasting ability. Although there is a lot of literature on GARCH-type models, no single GARCH-type model has been found to be superior to all others to model and forecast the time-varying systematic covariance risk.

Another technique for modelling and forecasting time-varying systematic covariance risk is based on the state space model estimated via Kalman Filter based approaches, which are recursive algorithms for estimating and forecasting unobserved time-varying systematic covariance (*beta*) risk. The most well-known Kalman Filter based approaches are the Random Coefficients (KFRC), Random Walk (KFRW), and Mean Reverting (KFMR) models. For example, Brooks et al. (1998) investigated the time-varying systematic covariance risk for both in-sample and out-of-sample procedures using the Kalman Filter based approaches for an Australian industry portfolio. Mergner and Bulla (2008) examined the KFRW and KFMR to determine the behaviour of the time-varying systematic covariance risk for 18 pan-European sectors in both the in-sample modelling and out-of-sample forecasting performance. Using UK industry sectors data from January 1969 to April 1998, Faff et al. (2000) employed the Kalman Filter based approaches to estimate the behaviour of the time-varying systematic covariance risk. Choudhry and Wu (2009) also investigated the weekly stock returns forecasts of 20 UK firms from January 1989 to December 2003 using the KFRW approach to evaluate the time-varying betas' accuracy in an out-of-sample forecasting procedure. Although there is much literature on the Kalman Filter based approaches, no single Kalman Filter model has been found to be superior to all others to model and forecast the time-varying systematic covariance risk.

The main purpose of this chapter is to confirm the instability of the systematic covariance (*beta*) risk in an emerging market, by comparing the modelling and forecasting abilities of the unconditional and conditional Two-Moment CAPMs. For the latter GARCH-type models such as GARCH and GJR-GARCH with normal and t (capturing heavy tails on returns) conditional distributions, and the Kalman Filter based approaches such as KFRC, KFRW and KFMR are used. The aim of this chapter is to compare the performance of these models, to contribute to the literature about the possible time-varying systematic covariance (*beta*) risk and possible approaches for capturing it. The comparison is made using weekly data, generated by 19 Turkish industry sector portfolios over the period from 1 August 2002 to 16 February 2012. In all cases the Istanbul Stock Exchange (ISE) All-Share index and the three-month Turkish Interbank Offer Rate (TRLIBOR) interest rate are used as a proxy for the market portfolio and risk-free rate, respectively. The modelling and forecasting abilities of models are evaluated using two different summaries of the errors, the Mean Square Error (MSE) and the Mean Absolute Error (MAE).

The rest of this chapter is outlined as follows. Section 4.2 presents the models to be compared and section 4.3 presents a description of the data. Section 4.4 presents the empirical results obtained from the model comparison, and section 4.5 presents further results from the best fitting model. Section 4.6 presents our conclusions.

4.2 Methodology

4.2.1 Linear Market Model

The Data Generating Process (DGP) of the Two-Moment CAPM is the Linear Market Model, which is described in section 2.1.2 where more theoretical details are provided. The Linear Market Model can be written as

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1i}(R_{mt} - R_{ft}) + \varepsilon_{it}, \tag{4.1}$$

where R_{it} and R_{mt} are the returns for industry sector i (i=1,...,19) and the ISE market portfolio at time t (t=1,...,T) respectively. R_{ft} is the risk-free rate at time t, and ε_{it} are the residuals with $\varepsilon_{it} \sim N(0, \sigma_i^2)$ and $E(\varepsilon_{it}\varepsilon_{kt})$=0, for $i \neq k$ and $E(\varepsilon_{it}\varepsilon_{i\,t+j})$=0 for $j > 0$. Here, the regression intercept, κ_i, and the regression slope, α_{1i} accounts for the systematic covariance (*proxy for* β_{im}) risk. These are constant over time. To show the link between the Two-Moment CAPM and the Linear Market Model (4.1), the systematic covariance (β_{im}) risk can be expressed as

$$\alpha_{1i} = \beta_{im} = \frac{Cov(R_{it} - R_{ft}, R_{mt} - R_{ft})}{Var(R_{mt} - R_{ft})}, \tag{4.2}$$

where the proof has been displayed in section 2.1.2. Note that the value of κ_i is expected to be zero in the Two-Moment CAPM, because the risk-free rate (R_{ft}) is subtracted before estimation (see e.g. Campbell et al. (1997), Faff et al. (2000), Mergner and Bulla (2008) and Choudhry and Wu (2009)). Hence, κ_i will be assumed to be zero for the rest of this chapter. Note that the parameter estimates of the Linear Market Model are obtained using the *lm* function from the *stats* package in the *R* software.

4.2.2 GARCH-type Models

To estimate and asses the instability of the time-varying systematic covariance (*beta*) risk, the time-varying Linear Market Model (consistent with a conditional Two-Moment CAPM) can be extended to allow β_{im} to evolve over time. Here, the estimation of β_{imt} is based on GARCH-type models, such as GARCH and GJR-GARCH with conditional distributions that are normal or t with the latter allowing for heavy tails. These models represent the time-varying systematic covariance risk indirectly by estimating the conditional variance of, and assuming a constant correlation between, the industry sector i and ISE market portfolio excess returns. In accordance with equation (3.56), the simplest model considered is the GARCH(1,1) model which is given by

$$\sigma_{it}^2 = \omega_i + \psi_{1i}(R_{i\,t-1} - R_{f\,t-1})^2 + \theta_{1i}\sigma_{i\,t-1}^2, \quad t = 2, \ldots, T, \tag{4.3}$$

in which the variance for industry sector i at time t, σ_{it}^2, (with $i = m$ for ISE market) depends on a constant, the excess returns and the conditional variance lagged by one time period. The parameters are restricted to be $\omega_i > 0$, $\psi_{1i} \geq 0$ and $\theta_{1i} \geq 0$ to ensure positive conditional variance at every time t. To ensure stationarity, it is required that $\psi_{1i} + \theta_{1i} < 1$.
For the GJR-GARCH model (equation (3.59))

$$\sigma_{it}^2 = \omega_i + \psi_{1i}(R_{i\,t-1} - R_{f\,t-1})^2 + \zeta_{1i}I_{t-1}(R_{i\,t-1} - R_{f\,t-1})^2 + \theta_{1i}\sigma_{i\,t-1}^2, \qquad (4.4)$$

for $t = 2, \ldots, T$, where I_{t-1} denotes an indicator function, taking a value of 1 if $(R_{i\,t-1} - R_{f\,t-1}) \leq 0$ and 0 otherwise.

Note that the conditional variance (equations (4.3) and (4.4)) starts at $t = 2$, since we do not know R_{i0} and R_{f0}. At $t = 1$, $\sigma_{i1}^2 = \frac{1}{T}\sum_{t=1}^{T}(R_{it} - R_{ft})^2$ is commonly set in the literature (e.g. Christoffersen (2003) and Danielsson (2011)) and the parameter estimation algorithm for GARCH-type models is briefly outlined in section 3.4.2. Parameter estimation for these models was implemented in the *rugarch* package (Ghalanos (2013)) in the *R* software.

The time-varying systematic covariance (*beta*) risk obtained from GARCH-type models is often expressed in the form

$$\beta_{imt} = \frac{Cov(R_{it} - R_{ft}, R_{mt} - R_{ft})}{Var(R_{mt} - R_{ft})} = \rho_{im}\frac{\sqrt{\sigma_{it}^2}}{\sqrt{\sigma_{mt}^2}}, \qquad (4.5)$$

where σ_{it}^2 and σ_{mt}^2 are the conditional variances of the industry sector i and ISE market excess returns at time t. Here, $\sigma_{imt} = Cov(R_{it} - R_{ft}, R_{mt} - R_{ft}) = \rho_{im}\sqrt{\sigma_{it}^2\sigma_{mt}^2}$ is the covariance between industry sector i and ISE market at time t, and for computational simplicity, ρ_{im} is assumed to be a time-invariant correlation coefficient between $R_{it} - R_{ft}$ and $R_{mt} - R_{ft}$. This is in accordance with the Constant Conditional Correlation GARCH (CCC-GARCH) model, introduced by Bollerslev (1990), in which the conditional variances of $R_{it} - R_{ft}$ and $R_{mt} - R_{ft}$ follow univariate GARCH-type models (equations (4.3) and (4.4)). Note that the constant correlation ρ_{im} is estimated using the *cor* function from the *stats* package in the *R* software.

4.2.3 Kalman Filter Based Models

The final set of models to be compared are the time-varying Linear Market Models, where the systematic covariance (β_{imt}) risk evolves over time via a Kalman Filter type model, such as Random Coefficient (KFRC), Random Walk (KFRW) and Mean Reverting (KFMR) as outlined in section 3.3. Model (4.1) with κ_i treated as zero is the observation equation of the state space model, and is expressed as follows.

$$R_{it} - R_{ft} = \alpha_{1it}(R_{mt} - R_{ft}) + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, H_i). \tag{4.6}$$

In accordance with equations (3.46, 3.47 and 3.48), the state equation can be expressed as any of the following three options:

KFRC

$$\alpha_{1it} = \bar{\alpha}_{1i} + w_{it}, \qquad w_{it} \sim N(0, Q_i), \tag{4.7}$$

where $\bar{\alpha}_{1i} = \frac{1}{T}\sum_{t=1}^{T} \alpha_{1it}$. This model suggests that shocks to the time-varying systematic covariance risk have no persistence from period to period.

KFRW

$$\alpha_{1it} = \alpha_{1i\,t-1} + w_{it}, \qquad w_{it} \sim N(0, Q_i). \tag{4.8}$$

This model suggests that shocks to the time-varying systematic covariance risk persist in the future.

KFMR

$$\alpha_{1it} = \bar{\alpha}_{1i} + \phi_i\left(\alpha_{1i\,t-1} - \bar{\alpha}_{1i}\right) + w_{it}, \qquad w_{it} \sim N(0, Q_i). \tag{4.9}$$

This model suggests that shocks to the time-varying systematic covariance risk have some persistence, but the coefficients return to their mean values (Faff et al. (2000)). These three models are finished with the prior specification

$$\alpha_{1i0} \sim N(\mu_{\alpha_{1i}}, \Sigma_{\alpha_{1i}}). \tag{4.10}$$

Here, $\mu_{\alpha_{1i}}$ and $\Sigma_{\alpha_{1i}}$ for α_{1i0} are set to the maximum likelihood estimates from the Linear Market Model. To show the link between the conditional (*time-varying*) Two-Moment CAPM and the time-varying Linear Market Model (4.6), the time-varying systematic covariance (β_{imt}) risk can be expressed as

$$\alpha_{1it} = \beta_{imt}, \tag{4.11}$$

which is an extension of the result proved in section 2.1.2. Note that software to implement the three Kalman Filter models is not generally available, and was written as part of this PhD thesis. The algorithm used is a modified version of that described in Shumway and Stoffer (2006), and is summarised in Appendix A of this thesis.

4.3 Data Description

The data for this study are weekly returns from 1 August 2002 to 16 February 2012 across 19 Istanbul Stock Exchange (ISE) industry sector indices maintained by the Borsa Istanbul (BIST) AS. This was founded on December 30, 2012 and combined the former Istanbul Stock Exchange (ISE), the Istanbul Gold Exchange (IAB) and the Derivatives Exchange of Turkey (VOB). The data were obtained from the Thomson Reuters Financial Datastream database provided by the University of Glasgow, UK on February 17, 2012. Table 4.1 represents an overview of all 19 industry sectors and their abbreviations utilized in this study. The main criteria that were used to select the industry sectors were: (1) industry sector classification based on predetermined criteria in the ISE database; and (2) continuous listing from 1 August 2002 to 16 February 2012. The market proxy for this study is the ISE National-All Share Index (ISE) capturing all National Market companies except investment trusts.

The one-week returns for all 19 industry sectors and the ISE market portfolio were obtained from the first difference in the logarithm of Wednesday's closing

Table 4.1: ISE industry sector classification.

Abbreviation	Industry Sector
Bank	Banks
Basic	Basic Materials
Chemical	Chemical & Petroleum & Plastic
Electricity	Electricity
Food	Food & Beverage
Holding	Holding & Investment
Info Tech	Information Technology
Insurance	Insurance
Investment	Investment Trusts
Leasing	Leasing & Factoring
Metal	Metal Goods & Machinery
Mineral	Non-Metal Mineral Products
REIT	Real Estate Investment Trusts
TeleCom	Telecommunications
Textile	Textile & Leather
Tourism	Tourism
Transport	Transportation
Wood	Wood & Chapter & Print
Wholesale	Wholesale & Retail Trade

Notes: Further details about ISE industry sector portfolios which are maintained by the Borsa Istanbul A.S. are available at http://borsaistanbul.com/en.

price expressed in Turkish liras as follows

$$R_{it} = \log(P_{it}) - \log(P_{i\,t-1}), \tag{4.12}$$

for $t = 2, \ldots, T$ and $i = 0, 1, \ldots, 19$, where $i = 0$ refers to the ISE market portfolio (R_{mt}). Here, P_{it} and $P_{i\,t-1}$ denote Wednesday's closing price in weeks t and $t-1$, respectively. The simple return r_{it}, meaning the gain or loss on an investment over the period $t-1$ to t, expressed as a proportion of the original investment, is represented as

$$\begin{aligned} r_{it} &= \frac{P_{it} - P_{i\,t-1}}{P_{i\,t-1}} = \frac{P_{it}}{P_{i\,t-1}} - 1, \\ 1 + r_{it} &= \frac{P_{it}}{P_{i\,t-1}}. \end{aligned} \tag{4.13}$$

Taking the logs (natural logarithm) of both sides of equation (4.13),

$$\log(1 + r_{it}) = \log\left(\frac{P_{it}}{P_{i\,t-1}}\right) = \log(P_{it}) - \log(P_{i\,t-1}). \tag{4.14}$$

When the one period simple return r_{it} is small, we can use a first-order Taylor expansion to approximate the simple return r_{it} and obtain

$$\log(1 + r_{it}) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{r_{it}^n}{n} \approx r_{it}, \quad -1 < r_{it} \leq 1. \tag{4.15}$$

Rewriting equation (4.14),

$$r_{it} \approx R_{it} = \log(P_{it}) - \log(P_{i\,t-1}). \tag{4.16}$$

The three-month Turkish Interbank Offered Rate (TRLIBOR) interest rate served as a proxy for the risk-free rate. As the TRLIBOR yields ($TRLIBOR_t$) are in percentage per annum, they can be converted to a weekly rate of return as follows (Mergner (2009)).

$$R_{ft} = \left(1 + \frac{TRLIBOR_t}{100}\right)^{1/52} - 1. \tag{4.17}$$

Table 4.2 shows descriptive statistics for the returns on the ISE market portfolio and the 19 industry sectors. The table details some key points. The mean return on the weekly ISE market portfolio is approximately 0.0036, with a standard deviation of about 0.0428. The range of mean weekly returns varies from 0.0008 for Electricity to 0.0046 for Basic, meaning that Basic generated greater financial gain on investment than Electricity during this period. The mean return on 13 out of 19 industry sectors is less than the mean risk-free rate (TRLIBOR), which represents the minimum return an investor theoretically expects for any investment, suggesting that investors would prefer not to invest in these industry sectors during this period. The highest unconditional volatility (*standard deviation*) is that for Tourism (0.0625), while the lowest one is that of Mineral (0.0359). Therefore, Tourism is deemed the riskiest sector when risk is measured by unconditional volatility.

Table 4.2: Descriptive statistics of weekly returns.

Industry	Mean	Std.Dev.[a]	Skewness	Kurtosis	JB^b	$LB(22)^c$
ISE	0.0036	0.0428	-0.10	7.10	354.21*	54.23*
Bank	0.0044	0.0555	0.36	6.80	314.66*	38.69*
Basic	0.0046	0.0554	-0.49	6.32	252.56*	166.35*
Chemical	0.0033	0.0446	-0.16	7.49	425.79*	61.80*
Electricity	0.0008	0.0534	-0.28	9.31	842.06*	82.92*
Food	0.0040	0.0404	-0.36	4.53	60.71*	52.15*
Holding	0.0024	0.0496	-0.28	6.28	232.29*	88.76*
Info Tech	0.0017	0.0475	-0.08	7.85	495.90*	47.50*
Insurance	0.0038	0.0571	-0.19	6.63	280.87*	54.04*
Investment	0.0027	0.0470	1.05	17.04	4223.66*	22.18
Leasing	0.0032	0.0587	-0.18	7.92	512.53*	19.35
Metal	0.0025	0.0460	-0.48	7.04	362.62*	73.55*
Mineral	0.0034	0.0359	-0.36	7.35	408.11*	49.02*
REIT	0.0027	0.0490	-0.05	10.85	1294.14*	31.17
TeleCom	0.0032	0.0533	0.10	5.16	99.36*	48.76*
Textile	0.0022	0.0415	-0.89	6.47	319.61*	48.54*
Tourism	0.0023	0.0625	0.50	11.01	1367.75*	74.39*
Transport	0.0028	0.0566	-0.49	5.68	171.51*	141.67*
Wood	0.0017	0.0466	-0.34	4.71	71.49*	37.40*
Wholesale	0.0043	0.0436	-0.28	19.88	5974.64*	104.43*
risk-free rate	0.0033	0.0018	1.26	4.14	161.31*	8051.22*

Notes: The portfolio has 498 observations for the weekly returns for each of the 19 industry sector portfolios. [a] Std.Dev. is the standard deviation. [b] *JB* is the Jarque-Bera statistic for testing the normality. *JB* follows χ^2 with 2 degrees of freedom so the critical value at the 5% level is 5.99. [c] *LB(22)* is the Ljung-Box test statistic for the null hypothesis of no autocorrelation in the squared returns up to order $\sqrt{498} \approx 22$. *LB* statistic follows χ^2 with 22 degrees of freedom so the critical value at the 5% level is 33.92. * means the appropriate null hypothesis is rejected at the 5% significance level.

The return distributions of all 19 industry sectors, except for Bank, Investment, TeleCom and Tourism and the ISE market portfolio exhibit negative skewness, while the risk-free rate returns exhibit positive skewness. Negative skewness means that there are frequent small increases and a few extreme drops in returns, while positive skewness means that there are frequent small drops and a few extreme increases in returns (see Algieri (2012)). The range of skewness varies between -0.89 for Textile and 1.05 for Investment. This suggests that Investment experiences frequent small drops and a few extreme increases, while Textile reports increases and a few extreme drops in terms of investment returns.

The return distributions of all 19 industry sectors, ISE market portfolio and risk-free rate are leptokurtic, meaning that the market has fatter tails than the normal distribution (which has kurtosis=3) and more chance of extreme outcomes. The range of kurtosis varies between 4.53 (Food) and 19.88 (Wholesale). This suggests that Wholesale has more chance of extreme financial losses or gains than the other sectors. The normality of each industry sector, ISE market portfolio and risk-free rate is also rejected at the 5% significance level using the Jarque-Bera (JB) test (see details in section 3.5.2.1) which is likely to be due to the substantial skewness and kurtosis observed from Table 4.2. To test the autocorrelation for the squared returns (proxies for volatilities) of the 19 industry sectors, the ISE market portfolio and the risk-free rate, the Ljung-Box ($LB(22)$) test (see details in section 3.5.2.1) is used in this thesis. According to the Ljung-Box ($LB(22)$) test, the null hypothesis of no autocorrelation for the squared returns is rejected at the 5% significance level for 16 out of the 19 industry sectors and the ISE market portfolio and the risk-free rate, meaning that there exists a statistically significant autocorrelation for the squared returns. These results provide strong evidence for the predictability of the volatility of the 16 industry sectors, the ISE market portfolio and the risk-free rate (Christoffersen (2003)). These results provide a positive effect over the performance of all models while predicting the time-varying volatility in the next section.

Overall, the main features of these data are the positive mean, relatively high volatility, asymmetry (left-skew and right-skew), and leptokurtosis (fat tails). These findings match the most common features of emerging market studies. (e.g. Harvey (1995) and next chapter of this thesis). This justifies the consideration of models that capture time-varying variance and covariance, rather than simply the Two-Moment CAPM for each industry sector portfolio.

Figure 4.1 displays the time series plots of returns on the ISE market and Bank, Chemical, TeleCom, and Tourism industry sectors, respectively. To save space the remaining industry sectors are displayed alphabetically and are a smaller size in Figures 4.2 and 4.3. The figures reveal some key points. Extreme events appear in 2002, resulting from the effect of the Turkish Stock Market Crash of 2001, and in October 2008, when a financial crisis began its global spread. It can be clearly seen that all industry sectors and the ISE market displayed even more volatility and extreme values in those weeks. These findings match the behaviour

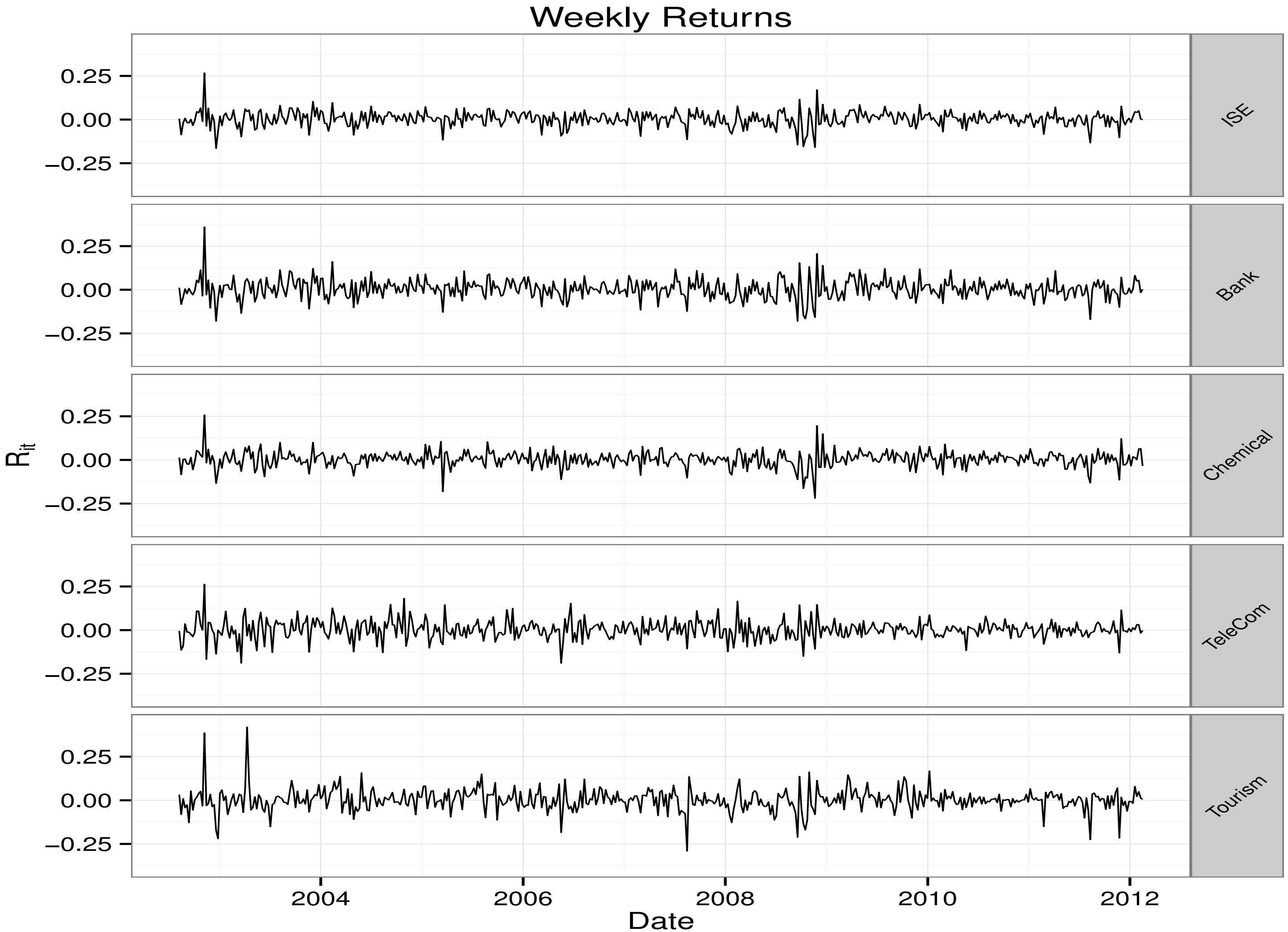


Figure 4.1: The time series plot of weekly returns on the ISE market and 4 industry sectors.

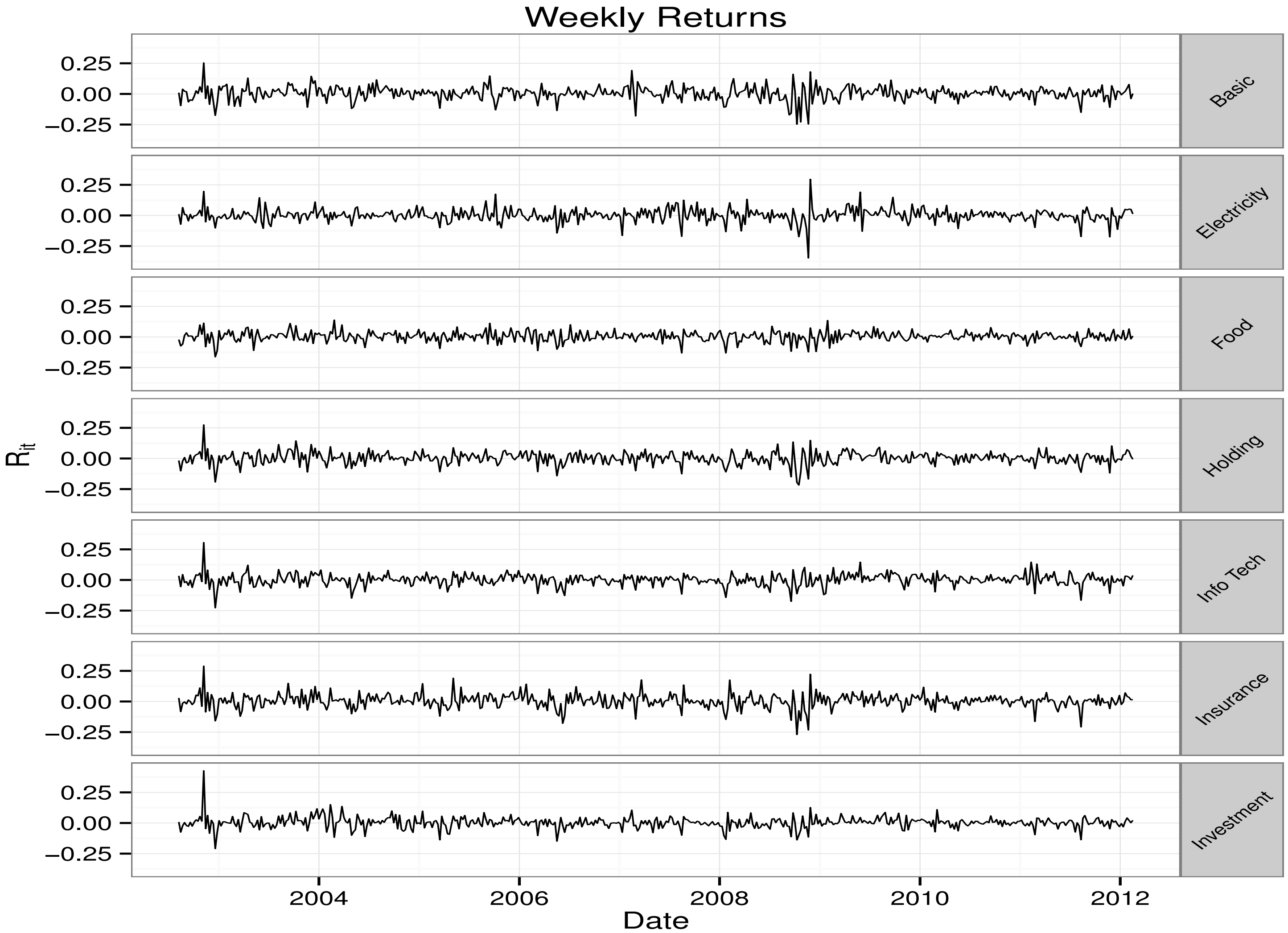


Figure 4.2: The time series plot of weekly returns on 7 industry sectors.

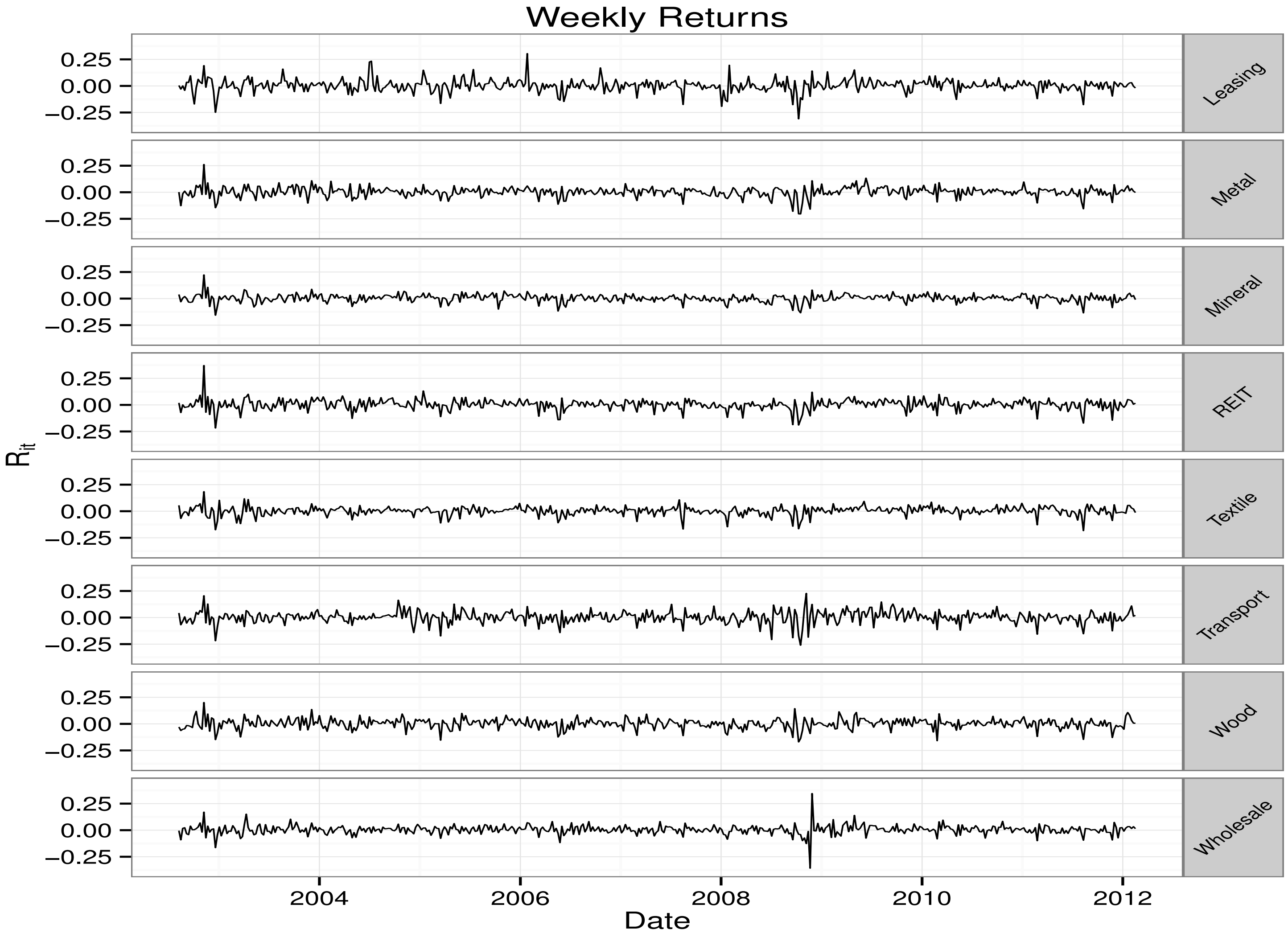


Figure 4.3: The time series plot of weekly returns on 8 industry sectors.

described in previous emerging market studies such as Harvey (1995) and that presented in the next chapter of this thesis. Even though these extreme values can affect the modelling and forecasting of asset returns, they are included in the dataset, since extreme behaviour is an inherent characteristic of financial markets (Ranaldo and Favre (2005)).

4.4 Comparison of Models

4.4.1 In-sample Model Fit

This section presents a comparison of the in-sample model fit performance for the Linear Market Model, GARCH-type models (GARCH and GJR-GARCH with conditional distributions normal and t) and Kalman Filter based approaches (KFRC, KFRW and KFMR). Note that while the GARCH-type models calculate the time-varying betas indirectly by estimating the conditional variance of (assuming a constant correlation) the industry sector i and ISE market portfolio excess returns, the time-varying betas are derived by Kalman Filter based approaches directly, as outlined in section 4.2. The comparison of model performance is in terms of two different measures of error, which are the Mean Absolute Error (MAE) and Mean Square Error (MSE), as outlined in section 3.5.1. The MAE and MSE values across all industry sectors for all modelling techniques in the in-sample procedure are presented in Tables 4.3 and 4.4, respectively.

A comparison of the different modelling techniques in the in-sample procedure results demonstrate overwhelming support for the Kalman Filter based approaches. Even though the average MAE and MSE of the KFRC model, equalling 1.825 and 6.914, respectively, are slightly lower than the average MAE (1.834) and MSE (6.961) for the KFMR, the MAE (MSE) values of the KFMR are equivalent to or less than those of the KFRC in 10 (9) out of 19 industry sectors. In addition, the KFMR is the preferable model as it is a general form that encompasses the KFRC ($\phi_i = 0$) and KFRW ($\phi_i = 1$) models as special cases.

It can be seen clearly that the average MAE and MSE values of the Kalman Filter based approaches are lower than those for the GARCH-type models re-

Table 4.3: MAE ($\times 10^2$) of in-sample model fit.

Industry	LMM	G-n	G-t	GJR-n	GJR-t	KFRW	KFRC	KFMR
Bank	1.182	1.190	1.196	1.207	1.305	1.182	**0.900**	**0.900**
Basic	2.470	2.475	2.500	2.484	2.513	2.448	**2.164**	2.178
Chemical	1.889	1.870	1.881	1.874	1.866	1.729	**1.590**	1.666
Electricity	2.860	2.801	2.799	2.816	2.847	2.722	**2.328**	2.332
Food	2.380	2.359	2.359	2.353	2.357	2.350	**1.981**	**1.981**
Holding	1.213	1.218	1.237	1.222	1.306	1.152	**1.002**	1.026
Info Tech	2.240	2.225	2.237	2.225	2.256	2.187	1.989	**1.977**
Insurance	2.212	2.223	2.224	2.240	2.447	2.172	**1.809**	**1.809**
Investment	2.159	2.168	2.159	2.168	2.255	2.036	**1.758**	**1.758**
Leasing	3.013	3.025	3.039	3.049	3.140	2.979	**2.507**	**2.507**
Metal	1.685	1.711	1.673	1.691	1.789	1.669	1.411	**1.408**
Mineral	1.573	1.557	1.570	1.557	1.558	1.531	**1.338**	1.341
REIT	1.949	1.904	1.933	1.899	2.065	1.723	**1.504**	1.546
TeleCom	2.979	2.973	2.971	2.991	3.005	2.942	**2.378**	**2.378**
Textile	2.094	2.080	2.071	2.082	2.144	1.938	**1.569**	1.585
Tourism	3.397	3.423	3.419	3.422	3.561	3.361	2.774	**2.725**
Transport	2.909	2.909	2.887	2.909	2.978	2.867	2.477	**2.476**
Wood	2.162	2.169	2.142	2.099	2.127	2.040	**1.566**	1.596
Wholesale	2.085	2.081	2.079	2.048	2.102	1.910	**1.638**	1.659
Average	2.234	2.230	2.230	2.228	2.296	2.155	**1.825**	1.834

Notes: The abbreviations of models are: LMM: Linear Market Model, G-n: GARCH-normal distribution, G-t: GARCH-t distribution, GJR-n: GJR-GARCH-normal distribution, GJR-t: GJR-GARCH-t distribution, KFRW: Kalman Filter Random Walk, KFRC: Kalman Filter Random Coefficient, KFMR: Kalman Filter Mean Reverting. **Bold** displays the best model for each industry sector in terms of lowest MAE.

gardless of the choice of conditional distributions in the case of each industry sector, meaning that GARCH-type models perform relatively poorly. Comparing the GARCH-type models, the average MAE (2.228) of the GJR-GARCH-normal is close to the average MAE (2.230) of the GARCH-normal and GARCH-t models. The average MAE of the GJR-GARCH-t is higher than that for the other GARCH-type models. Within the GARCH-type models, the lowest average MSE is for the GJR-GARCH-normal at 10.249, while the highest average MSE for the GJR-GARCH-t is 10.831. This suggests that improved modelling cannot be clearly observed when using t distributed errors compared with normal distributed errors.

In addition, the modelling performance of the Linear Market Model is clearly worse than the time-varying Linear Market Model via Kalman Filter based approaches, but similar to the GARCH-type models. This suggests that the extra

Table 4.4: MSE ($\times 10^4$) of in-sample model fit.

Industry	LMM	G-n	G-*t*	GJR-n	GJR-*t*	KFRW	KFRC	KFMR
Bank	2.389	2.436	2.464	2.483	3.019	2.389	**1.371**	**1.371**
Basic	11.503	11.505	11.880	11.419	11.458	11.216	**8.792**	8.965
Chemical	5.982	5.783	5.875	5.844	5.760	4.992	**4.247**	4.614
Electricity	16.139	15.841	16.032	15.677	16.183	14.687	**10.685**	10.718
Food	9.844	9.756	9.766	9.669	9.720	9.601	**6.768**	6.832
Holding	2.656	2.685	2.804	2.715	3.144	2.362	**1.768**	1.864
Info Tech	9.659	9.549	9.640	9.505	9.723	9.214	7.687	**7.642**
Insurance	9.808	10.118	10.010	10.414	12.101	9.424	**6.672**	**6.672**
Investment	9.703	9.733	9.781	9.702	10.226	8.328	**6.353**	6.376
Leasing	19.323	19.674	19.449	19.700	20.552	18.820	**13.682**	**13.682**
Metal	5.167	5.611	5.228	5.313	6.122	5.083	3.630	**3.618**
Mineral	4.295	4.292	4.288	4.279	4.318	4.039	**3.140**	3.167
REIT	6.577	6.367	6.450	6.321	7.358	5.078	**3.925**	4.114
TeleCom	16.193	15.913	16.045	16.029	16.222	15.633	**10.356**	**10.356**
Textile	8.121	7.954	7.980	7.965	8.636	6.821	**4.622**	4.695
Tourism	22.692	23.057	22.931	22.871	24.554	22.170	15.311	**15.051**
Transport	17.065	17.497	17.429	17.433	18.384	16.736	12.420	**12.415**
Wood	8.513	8.573	8.311	8.316	8.591	7.416	**4.469**	4.654
Wholesale	9.749	9.146	9.392	9.068	9.724	7.236	5.459	**5.445**
Average	10.283	10.289	10.303	10.249	10.831	9.539	**6.914**	6.961

Notes: The abbreviations of models are: LMM: Linear Market Model, G-n: GARCH-normal distribution, G-*t*: GARCH-*t* distribution, GJR-n: GJR-GARCH-normal distribution, GJR-*t*: GJR-GARCH-*t* distribution, KFRW: Kalman Filter Random Walk, KFRC: Kalman Filter Random Coefficient, KFMR: Kalman Filter Mean Reverting. **Bold** displays the best model for each industry sector in terms of lowest MSE.

complexity of the GARCH-type models does not lead to improved modelling performance compared with the Linear Market Model. This is surprising at first sight, but may be due to the constant correlation assumption implicit in the computation of the time-varying betas from the GARCH-type models. In contrast, the Kalman Filter based approaches offer a substantial improvement on the Linear Market Model. To summarise, of the different modelling techniques used for in-sample procedure, the KFMR seems to be best qualified to model the weekly time-varying systematic covariance (*beta*) risk for the 19 industry sector portfolios in the CAPM.

4.4.2 Out-of-sample Forecasting

This section presents a comparison of the forecasting performance of the same models, using an out-of-sample procedure which allows us to evaluate the models predictive performance. Recently (e.g. Tsay (2005) and Mergner and Bulla (2008)) a rolling window technique has been used to undertake an out-of-sample performance comparison, and this is the approach adopted here. Firstly, the length of the rolling window needs to be defined, and Tsay (2005) suggests that the length should be $T/2$ for large data or $2T/3$, where T is the length of data, which is enough data to generate stable parameter estimates. The length of the rolling window used in this study is 5 years (260 weeks), which is approximately equal to $T/2$. All models are fitted to 5 years worth of data, and each is then used to predict beta one-week ahead (one-step ahead prediction). The 5 years worth of data are then rolled forward by one week and the process of one-step ahead prediction is repeated. This process is continued for 2 years (104 weeks) which is short enough to reflect current market conditions (Mergner and Bulla (2008)), and the MAE and MSE between these predicted excess returns and the actual excess returns are computed over these 104 values. The predictions are over the period from February 25, 2010 to February 16, 2012. Tables 4.5 and 4.6 present the MAE and MSE measures across all industry sectors respectively, for each modelling technique for the out-of-sample forecasting process.

A comparison of the different modelling techniques resulting from the out-of-sample rolling window results show overwhelming support for the Kalman Filter based approaches. The average MAE and MSE values of the KFRC model, equalling 1.532 and 4.344, respectively, are slightly lower than the average MAE (1.541) and MSE (4.391) of the KFMR; whereas, the MAE (MSE) values of the KFMR are equal to or less than those of the KFRC for 11 (12) out of the 19 industry sector porfolios. Moreover, the average MAE and MSE of the KFRW, equalling 1.769 and 5.765, are higher than those for the KFMR and KFRC. To sum up, the KFMR is again the preferable model as it is the general form of the other two and competitive with the KFRC.

Within the GARCH-type models, the average MAE and MSE values for the

Table 4.5: MAE ($\times 10^2$) of out-of-sample forecasts.

Industry	LMM	G-n	G-t	GJR-n	GJR-t	KFRW	KFRC	KFMR
Bank	1.108	1.146	1.128	1.190	1.167	1.101	**0.880**	**0.880**
Basic	1.847	1.692	1.685	1.688	1.674	1.628	1.458	**1.453**
Chemical	1.869	1.779	1.775	1.748	1.754	1.608	**1.580**	1.620
Electricity	2.361	2.162	2.163	2.124	2.256	2.139	**1.853**	1.854
Food	2.926	2.049	2.036	2.054	2.042	1.997	1.705	**1.690**
Holding	1.116	1.028	1.025	1.015	1.019	0.949	**0.867**	0.873
Info Tech	2.535	2.248	2.202	2.218	2.218	2.224	2.073	**2.060**
Insurance	1.620	1.476	1.524	1.488	1.593	1.464	**1.251**	**1.251**
Investment	2.234	1.719	1.722	1.707	1.787	1.693	**1.291**	**1.291**
Leasing	2.516	2.274	2.394	2.249	2.374	2.152	**1.808**	**1.808**
Metal	1.791	1.625	1.701	1.607	1.649	1.622	1.398	**1.391**
Mineral	2.041	1.549	1.514	1.525	1.523	1.546	**1.355**	1.356
REIT	1.765	1.733	1.834	1.790	1.780	1.580	**1.348**	1.382
TeleCom	2.450	1.920	1.922	1.935	1.923	1.882	**1.586**	**1.586**
Textile	2.577	2.205	2.230	2.284	2.339	2.053	**1.783**	1.796
Tourism	2.468	2.149	2.335	2.265	2.254	2.208	1.829	**1.810**
Transport	2.246	2.226	2.229	2.211	2.213	2.229	**1.961**	**1.961**
Wood	2.071	1.922	1.908	1.888	1.932	1.750	**1.374**	1.493
Wholesale	2.509	2.033	2.021	2.065	2.050	1.792	**1.715**	1.727
Average	2.108	1.839	1.860	1.845	1.871	1.769	**1.532**	1.541

Notes: The abbreviations of models are: LMM: Linear Market Model, G-n: GARCH-normal distribution, G-t: GARCH-t distribution, GJR-n: GJR-GARCH-normal distribution, GJR-t: GJR-GARCH-t distribution, KFRW: Kalman Filter Random Walk, KFRC: Kalman Filter Random Coefficients, KFMR: Kalman Filter Mean Reverting. **Bold** displays the best model for each industry sector in terms of lowest MAE.

GARCH-normal and GARCH-t are slightly lower than those of the GJR-GARCH-normal and GJR-GARCH-t. The average MAE (1.839) and MSE (6.216) of the GARCH-normal are the lowest from the GARCH-type models considered, while the highest average MAE and MSE are for the GJR-GARCH-t, equalling 1.871 and 6.446, respectively. This suggests that an improved forecasting performance is not clearly evident when using t rather than normal distributions with the GJR-GARCH model. Also, the GARCH-normal model outperforms all other GARCH-type models. The average MAE and MSE values of the Linear Market Model, equal to 2.108 and 8.013, respectively, are higher than those for any time-varying modelling technique. In particular, the results are substantially worse than those from the GARCH-type models, which was not the case for the in-sample model comparison. However, overall, the KFMR seems to be the best model for weekly time-varying systematic covariance (*beta*) risk for the 19 industry sector portfolios

Table 4.6: MSE ($\times 10^4$) of out-of-sample forecasts.

Industry	LMM	G-n	G-t	GJR-n	GJR-t	KFRW	KFRC	KFMR
Bank	2.038	2.228	2.115	2.636	2.481	1.968	**1.227**	**1.227**
Basic	5.759	4.835	4.893	4.835	4.776	4.509	3.816	**3.785**
Chemical	6.086	4.916	4.916	4.834	4.837	4.207	**3.969**	4.194
Electricity	9.433	8.805	8.353	8.481	9.474	8.389	6.552	**6.545**
Food	14.660	6.719	6.684	6.710	6.668	6.411	4.767	**4.747**
Holding	2.203	1.751	1.795	1.707	1.715	1.419	**1.204**	1.223
Info Tech	13.022	11.286	10.731	10.997	10.797	10.985	9.515	**9.364**
Insurance	4.749	4.282	4.424	4.451	4.939	3.943	**2.713**	**2.713**
Investment	9.199	5.307	5.279	5.327	5.765	5.073	**2.973**	**2.973**
Leasing	9.936	8.582	9.363	8.288	9.100	8.020	**5.934**	**5.934**
Metal	5.355	4.353	4.986	4.296	4.489	4.328	3.162	**3.147**
Mineral	6.704	3.740	3.643	3.603	3.591	3.763	**2.894**	2.899
REIT	5.186	5.140	5.889	5.441	5.213	4.114	**3.210**	3.308
TeleCom	10.667	6.644	6.770	6.671	6.727	6.382	**4.350**	**4.350**
Textile	11.032	8.667	8.465	9.350	9.714	6.992	**5.485**	5.548
Tourism	10.073	8.589	11.583	10.253	9.130	9.400	5.823	**5.810**
Transport	8.100	8.159	8.204	8.368	8.299	8.156	**6.333**	**6.333**
Wood	8.164	6.991	6.943	7.087	7.519	5.726	**3.178**	3.842
Wholesale	9.883	7.107	7.051	7.280	7.236	5.744	**5.428**	5.479
Average	8.013	6.216	6.426	6.348	6.446	5.765	**4.344**	4.391

Notes: The abbreviations of models are: LMM: Linear Market Model, G-n: GARCH-normal distribution, G-t: GARCH-t distribution, GJR-n: GJR-GARCH-normal distribution, GJR-t: GJR-GARCH-t distribution, KFRW: Kalman Filter Random Walk, KFRC: Kalman Filter Random Coefficients, KFMR: Kalman Filter Mean Reverting. **Bold** displays the best model for each industry sector in terms of lowest MSE.

in the CAPM.

4.5 Time-Varying Linear Market Model via KFMR

As the KFMR is the general form of the KFRW and KFRC models outlined in section 4.2.3, and because it is one of the best performing models in section 4.4, we examine it in greater detail here. The hyperparameter estimates of the KFMR model using equations (4.6) and (4.9) for 19 weekly Turkish industry sectors are presented in Table 4.7.

The parameter ϕ_i summarises the temporal autocorrelation in $\{\alpha_{1it}\}_{t=1}^{N}$ and will lie in the range 0 to 1 for a stationary series (see details in section 3.3). Here, estimated $\hat{\phi}_i$ is close to 0 for 5 (Bank, Insurance, Leasing, TeleCom and Transport) out of the 19 industry sectors, so that the time-varying systematic

Table 4.7: Time-varying Linear Market Model hyperparameter estimates (standard errors) via KFMR.

Industry	$\hat{Q}_i \times 100$	$\hat{H}_i \times 100$	$\hat{\phi}_i$
Bank	4.144	0.017	0.000
	(1.113)	*(0.002)*	*(0.000)*
Basic	5.817	0.099	0.611
	(4.891)	*(0.009)*	*(0.373)*
Chemical	1.462	0.050	0.851
	(0.789)	*(0.004)*	*(0.524)*
Electricity	15.217	0.124	0.410
	(5.339)	*(0.010)*	*(0.121)*
Food	9.168	0.079	0.457
	(3.897)	*(0.006)*	*(0.168)*
Holding	1.843	0.021	0.630
	(1.132)	*(0.002)*	*(0.374)*
Info Tech	3.966	0.083	0.623
	(2.444)	*(0.006)*	*(0.311)*
Insurance	12.361	0.078	0.000
	(4.841)	*(0.007)*	*(0.000)*
Investment	9.874	0.074	0.234
	(3.824)	*(0.006)*	*(0.098)*
Leasing	20.472	0.158	0.000
	(7.321)	*(0.013)*	*(0.000)*
Metal	5.275	0.042	0.220
	(1.779)	*(0.003)*	*(0.066)*
Mineral	3.676	0.036	0.284
	(1.826)	*(0.003)*	*(0.169)*
REIT	4.285	0.048	0.745
	(2.127)	*(0.004)*	*(0.347)*
TeleCom	23.673	0.125	0.000
	(7.809)	*(0.011)*	*(0.000)*
Textile	13.156	0.058	0.209
	(4.452)	*(0.005)*	*(0.098)*
Tourism	23.674	0.176	0.437
	(8.458)	*(0.015)*	*(0.113)*
Transport	16.145	0.141	0.037
	(5.642)	*(0.011)*	*(0.029)*
Wood	16.415	0.059	0.195
	(6.878)	*(0.007)*	*(0.110)*
Wholesale	11.485	0.065	0.224
	(2.798)	*(0.005)*	*(0.052)*

Notes: *Italic* numbers in parentheses denote the standard errors of the time-varying Linear Market Model hyperparameter estimates via KFMR.

Table 4.8: Time-varying Linear Market Model state parameter estimates (standard errors) via KFMR.

Industry	$\hat{\bar{\beta}}_{imt}$	Range ($\hat{\beta}_{imt}$)
Bank	1.265	(0.848;1.651)
	(0.026)	
Basic	0.978	(0.491;1.614)
	(0.049)	
Chemical	0.826	(0.477;1.293)
	(0.039)	
Electricity	0.748	(0.098;1.909)
	(0.042)	
Food	0.589	(-0.429;1.160)
	(0.027)	
Holding	1.090	(0.840;1.569)
	(0.028)	
Info Tech	0.812	(0.434;1.223)
	(0.036)	
Insurance	1.079	(0.469;1.815)
	(0.045)	
Investment	0.720	(0.132;1.533)
	(0.030)	
Leasing	0.862	(-0.272;2.104)
	(0.049)	
Metal	0.911	(0.493;1.501)
	(0.028)	
Mineral	0.662	(0.347;0.969)
	(0.018)	
REIT	0.911	(0.376;1.428)
	(0.044)	
TeleCom	0.824	(-0.075;1.671)
	(0.044)	
Textile	0.650	(-0.126;1.324)
	(0.026)	
Tourism	0.870	(0.197;2.730)
	(0.060)	
Transport	0.892	(0.154;1.743)
	(0.047)	
Wood	0.852	(-0.311;1.711)
	(0.036)	
Wholesale	0.625	(-0.152;1.969)
	(0.025)	

Notes: Range displays the range of $\hat{\beta}_{imt} = \hat{\alpha}_{1it}$. *Italic* numbers in parentheses denote the standard errors of the time-varying Linear Market Model state parameter estimates via KFMR.

covariance (*beta*) series of the KFMR becomes similar to the KFRC. Therefore the MAE and MSE of the KFMR model are close to those of the KFRC model for these industry sectors (see in Tables 4.3 and 4.4). The time-varying systematic covariance (*beta*) series of the KFMR, on the other hand, becomes similar to the KFRW when ϕ_i is close to 1, although the highest estimated value for $\hat{\phi}_i$ in the data is 0.851 with a standard error of only 0.524 for Chemical.

The estimated values of $\hat{Q}_i$ are much higher than those of $\hat{H}_i$ for all 19 industry sectors, meaning that the state variance captures the volatility of the industry sector's excess returns more than the observation variance. In addition, Tourism has the highest estimated $\hat{Q}_i$ (23.674) and $\hat{H}_i$ (0.176) of all industry sectors, and it is notable that this sector also had the maximum unconditional volatility shown in Table 4.2.

Table 4.8 presents mean and range for the time-varying systematic covariance (*beta*) parameter estimates of the KFMR (equations (4.6) and (4.9)) for all 19 Turkish industry sectors. The mean of the time-varying systematic covariance $\hat{\beta}_{imt}$ for all of the 19 industry sectors is positive, and is close to 1, with a standard error close to 0.04. Note that a systematic covariance ($\hat{\beta}_{im}$) value of 1 means that the industry sector moves in step with the ISE market portfolio. A value of $\hat{\beta}_{im}$ in the range from 0 to 1 means that the industry sector is less volatile than the ISE market portfolio, whereas a $\hat{\beta}_{im}$ value greater than 1 indicates that the sector is more volatile than the ISE market portfolio. In addition, a value ($\hat{\beta}_{im}$) of 0 would mean that the industry sector is not correlated to the ISE market portfolio, and a negative ($\hat{\beta}_{im}$) value would mean that the industry sector moves in the opposite direction with relation to the ISE market portfolio.

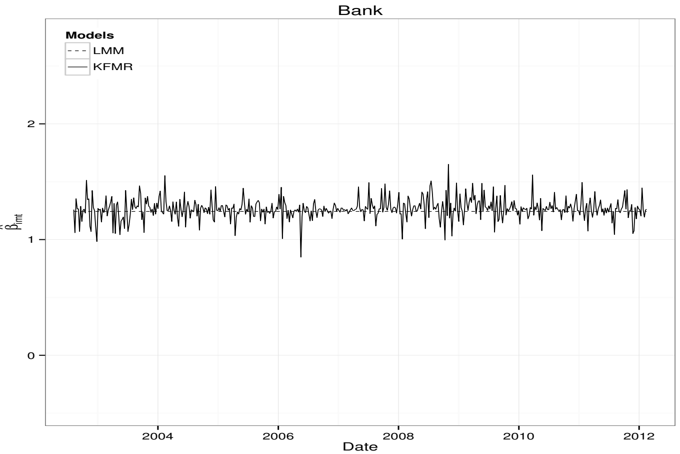
The final column of Table 4.8 shows the range of estimated beta values across the time period. The wider the range of values, the less consistent is the relationship between excess returns in that sector and the market as a whole. For Tourism, the KFMR provides a wider range of time-varying systematic covariance ($\hat{\beta}_{imt}$) than the other sectors, while Mineral has a narrower range than the other sectors. As can be seen from Table 4.2, Tourism is also the sector with the highest, and Mineral the sector with the lowest, unconditional volatility. Also, negative $\hat{\beta}_{imt}$ values appear in the Food, Leasing, TeleCom, Textile, Wood and

Wholesale sectors, meaning that, from time to time, these sectors move in the opposite direction to the ISE market portfolio.

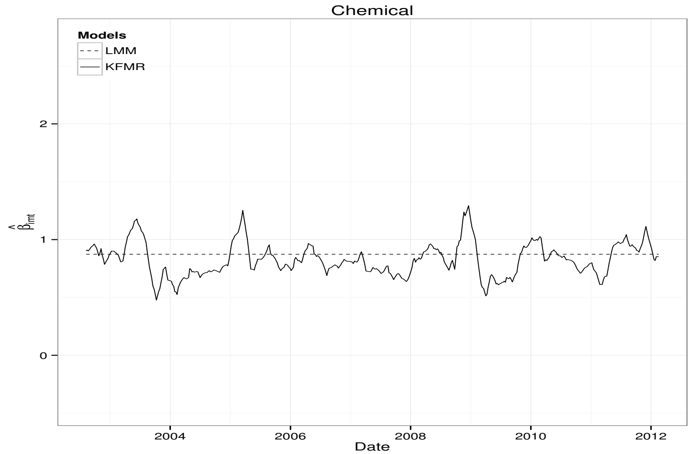
Figure 4.4 also displays the time-varying systematic covariance ($\hat{\beta}_{imt}$) risk series of Bank, Chemical, TeleCom, and Tourism industry sectors, respectively. These industry sectors have been chosen to illustrate the effects on the time-varying beta series of the speed parameter, ϕ_i, and the state variance, Q_i. To save space the remaining industry sectors are not displayed in this thesis. It can be clearly seen that the time-varying systematic covariance ($\hat{\alpha}_{1it} = \hat{\beta}_{imt}$) risk series fluctuates about the equivalent Linear Market Model (LMM) estimate ($\hat{\alpha}_{1i} = \hat{\beta}_{im}$) in all 4 industry sectors.

$\hat{\phi}_i$ controls the speed of fluctuation (temporal autocorrelation) in the series of time-varying beta values. The Chemical sector has the highest $\hat{\phi}_i$ value (0.851, close to 1 which is consistent with a random walk model), while $\hat{\phi}_i = 0.000$ (which is consistent with a random coefficient model) for both Bank and TeleCom, in the Tourism ($\hat{\phi}_i = 0.437$) somewhere between these extremes. The closer $\hat{\phi}_i$ is to 0, the less correlated successive $\hat{\beta}_{imt}$ values are and the more rapidly the series changes.

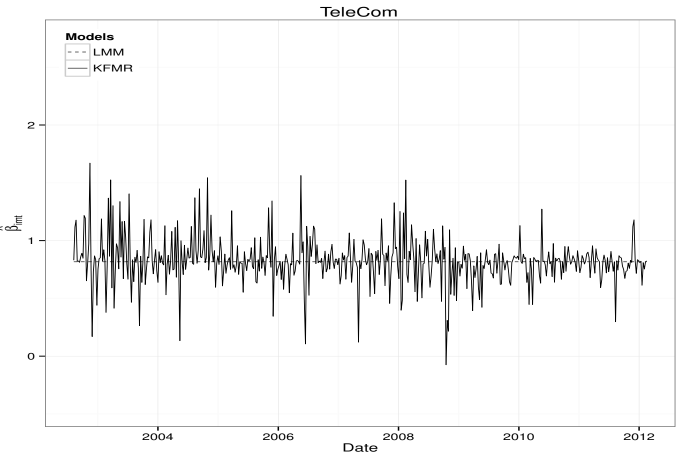
On the other hand, $\hat{Q}_i$ controls the amplitude of the overall fluctuation in the betas. Bank and Chemical have relatively low $\hat{Q}_i$ values, whereas TeleCom and Tourism have the highest values and it is clear that there is much larger variation in $\hat{\beta}_{imt}$ in those industry sectors.



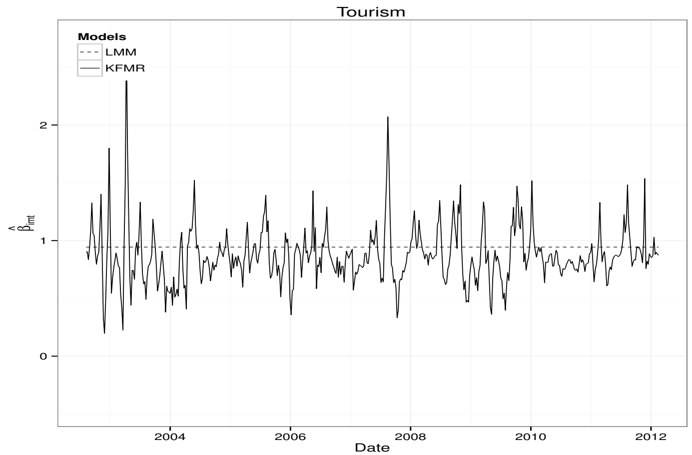
(a) Bank



(b) Chemical



(c) TeleCom



(d) Tourism

Figure 4.4: The estimated $\hat{\beta}_{imt} = \hat{\alpha}_{1it}$ plots of Bank, Chemical, Telecom, and Tourism industry sectors.

Table 4.9: Diagnostic test statistics for KFMR.

Industry	JB	$Het(166)$	$LB(22)$	Industry	JB	$Het(166)$	$LB(22)$
Bank	12.80*	0.97	18.83	Metal	404.69*	2.57*	25.12
	(0.002)	*(0.578)*	*(0.468)*		*(0.000)*	*(0.000)*	*(0.157)*
Basic	135.27*	0.69	18.49	Mineral	84.33*	0.85	37.23*
	(0.000)	*(0.991)*	*(0.490)*		*(0.000)*	*(0.852)*	*(0.007)*
Chemical	39.68*	1.07	23.88	REIT	42.56*	0.95	19.95
	(0.000)	*(0.332)*	*(0.201)*		*(0.000)*	*(0.629)*	*(0.398)*
Electricity	483.68*	1.34*	39.82*	TeleCom	72.13*	0.34	36.62*
	(0.000)	*(0.030)*	*(0.003)*		*(0.000)*	*(0.999)*	*(0.009)*
Food	51.07*	0.87	44.96*	Textile	242.60*	1.62*	27.12
	(0.000)	*(0.815)*	*(0.001)*		*(0.000)*	*(0.001)*	*(0.102)*
Holding	35.69*	1.39*	20.79	Tourism	1474.47*	0.44	12.02
	(0.000)	*(0.017)*	*(0.349)*		*(0.000)*	*(0.999)*	*(0.885)*
Info Tech	494.56*	2.20*	28.52	Transport	354.81*	1.04	14.79
	(0.000)	*(0.000)*	*(0.074)*		*(0.000)*	*(0.400)*	*(0.736)*
Insurance	942.97*	1.16	22.73	Wood	54.34*	1.21	50.95*
	(0.000)	*(0.171)*	*(0.249)*		*(0.000)*	*(0.110)*	*(0.000)*
Investment	332.06*	0.38	13.27	Wholesale	346.36*	2.20*	31.54*
	(0.000)	*(0.999)*	*(0.824)*		*(0.000)*	*(0.000)*	*(0.035)*
Leasing	1298.84*	0.53	39.86*				
	(0.000)	*(0.999)*	*(0.003)*				

Notes: *JB* is the Jarque-Bera test statistic for the null hypothesis of normally distributed standardised residuals. *JB* follows χ^2 with 2 degrees of freedom so the critical value at the 5% level is 5.99. *LB(22)* is the Ljung-Box test statistic for the null hypothesis of no autocorrelation in the standardised residuals up to order $\sqrt{498} \approx 22$. *LB* statistic follows χ^2 with 22-(m-1) degrees of freedom where m is the total number of estimated parameters. The relevant critical value at the 5% level is 30.14 (Harvey (1989)). *Het(166)* is the test statistic for the null hypothesis of no heteroskedasticity in the standardised residuals up to order $498/3 = 166$. *Het(166)* statistic follows $F_{(166,166)}$ distribution so the critical value at the 5% level is 1.29. * means the appropriate null hypothesis is rejected at the 5% significance level.

Table 4.9 presents the diagnostic test statistics for the residuals from KFMR model outlined in section 3.5.2.1. According to the Jarque-Bera (JB) test, the residuals are not normally distributed at the 5% significance level for all 19 industry sectors, implying that KFMR is poor in terms of non-normal errors. According to the H ($Het(166)$) test, the null hypothesis of no heteroskedasticity cannot be rejected in 13 out of the 19 industry sectors at the 5% significance level, implying that KFMR is adequate in terms of no heteroskedasticity for more than half of the 19 industry sectors. According to the Ljung-Box ($LB(22)$) test, the null hypothesis of no autocorrelation cannot be rejected at the 5% significance level for 12 out of the 19 industry sectors, meaning that KFMR is adequate in terms of no autocorrelation for more than half of the 19 industry sectors.

4.6 Conclusion

The focus of this chapter has been on confirming the instability of the systematic covariance (*beta*) risk in Turkish industry sector portfolios, by comparing the modelling and forecasting abilities of both unconditional and conditional Two-Moment CAPMs. In both cases the performance of the Linear Market Model is compared with that of the GARCH-type models and the Kalman Filter based approaches.

The performance of the different modelling techniques, when using the in-sample and out-of-sample procedures, were evaluated using MAE and MSE. The Linear Market Model is generally worse than the time-varying modelling techniques according to both criteria in the in-sample and out-of-sample procedures, but the GARCH-type models do not provide improved modelling performance compared with the Linear Market Model. In addition, within the GARCH-type models, an increase in performance can be clearly observed when assuming normal distributed errors compared with t distributed errors in both the in-sample and out-of-sample procedures. The model evaluation criteria clearly show that Kalman Filter based approaches provide a much better performance than any of the others in both the in-sample and out-of-sample procedures. In addition, within the Kalman Filter based approaches, the KFMR specification seems to be the best as it is the most flexible model.

The results confirm the instability of the systematic covariance (*beta*) risk in the Two-Moment CAPM found in the existing literature. In addition, this chapter confirms previous studies (e.g. Australia (Brooks et al. (1998)), and UK (Faff et al. (2000))) which found that Kalman Filter based approaches outperform GARCH-type models. In addition, both KFRC and KFMR models outperform KFRW, and are generally found to be superior to all other models to model and forecast the time-varying systematic covariance (*beta*) risk.

Note that in section 4.5, diagnostic procedures are discussed to check how far the assumptions of the state space model are satisfied. These assumptions are that the residuals are normally distributed, independent (no autocorrelation) and have constant variance (no heteroskedasticity). When these assumptions are violated, the performance of the state space model can be affected, though there are possible extensions of the state space model as discussed below.

The first assumption is that the distribution of the residuals is Gaussian. This assumption is violated in the modelling described in this chapter, which is a consequence of asymmetry and heavy tails due to unexpected increases and drops in asset returns, which may be caused by national events (e.g. wars or disasters), political events (e.g. elections) and economic factors (e.g. unemployment rate). The Gaussian distribution can be replaced by heavy-tailed distributions such as the t distribution, or a mixture of normals, or a general residual distribution (see Durbin and Koopman (2001)), or by an asymmetric distribution, such as a skewed-t distribution. In fact, Meinhold and Singpurwalla (1989) showed that using a t distribution for the residuals increases the robustness of the Kalman Filter against outliers.

The second assumption is that the variance of residuals is assumed to be constant. In finance, stock market time series data are subject to temporally non-constant fluctuations resulting from changing market conditions, therefore a constant variance assumption is often unrealistic. When the constant variance (homoskesdaticity) assumption is violated, researchers can use the stochastic volatility model, which allows them to capture a time-varying variance. This model is similar to the state space model, and various extensions of stochastic volatility models exist, allowing a combined GARCH-type and state space model, with not only Gaussian but also non-Gaussian residuals (see Durbin and Koopman (2001)).

The final assumption is that the residuals are independent (no autocorrelation). In this thesis, this assumption is invalid. One approach would be to allow the intercept term to vary over time via a random walk model within the Kalman Filter algorithm, in addition to the slope parameter β_{imt}.

Chapter 5

Is the Linear Market Model appropriate for Developed and Emerging Markets?

5.1 Introduction

In the financial literature the Two-Moment Capital Asset Pricing Model (CAPM) of Sharpe-Lintner-Mossin (1960s) is the most widely used framework for exploring systematic covariance (*beta*) risk. It depends on two restrictive assumptions, namely that asset returns are normally distributed and that the investor's utility function is quadratic (so, it can be expressed in terms of just the mean and variance of wealth). However, literature (e.g. Kraus and Litzenberger (1976), Fang and Lai (1997) and Hwang and Satchell (1999)) suggests that the Two-Moment CAPM may be misleading and insufficient to characterize asset returns, since returns on many assets are now believed to be non-normally distributed. These inadequacies of the Two-Moment CAPM have encouraged financial researchers to explore beyond the benchmark Linear Market Model.

One extension in this vein is to incorporate higher order moments into the Two-Moment CAPM. In the literature, the Higher-Moment CAPMs, namely the Three-Moment and Four-Moment CAPMs can capture the systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*) in financial data. The necessity for these is assessed by fitting Higher order Data Generating Processes (DGPs) to the financial data, namely the Quadratic and Cubic Market Models. For example, Barone-Adesi (1985), Fang and Lai (1997) and Hwang and Satchell (1999) proposed several formulations of Higher order DGPs for the purposes of illustrating the link between the Higher order DGPs and their equivalent Higher-Moment CAPMs and reducing the multicollinearity of the systematic risk measures in the Higher-Moment CAPMs.

Research by Hwang and Satchell (1999) assesses the appropriateness of Higher-Moment CAPMs via Higher order DGPs for monthly returns in emerging markets but not developed markets. They conclude that emerging markets are better explained by including additional systematic risk measures, such as *co-skewness* and *co-kurtosis* into the model. Fang and Lai (1997) also confirm that not only *co-skewness* but also *co-kurtosis* has an important role for explaining returns on the New York Stock Exchange (NYSE). Barone-Adesi (1985) proposes the Quadratic Market Model (consistent with the Three-Moment CAPM only capturing *co-skewness*) for security pricing. More recently, Hung (2007) suggests that the Quadratic Market Model is useful for explaining time-series weekly returns for developed markets.

Other financial economists have focused on confirming the instability of the systematic covariance (*beta*) risk (e.g. Brooks et al. (1998) and Faff et al. (2000)), by comparing the forecasting ability of the unconditional and conditional (*time-varying*) Two-Moment CAPMs. In the recent literature (e.g. Mergner and Bulla (2008) and Choudhry and Wu (2009)) comparisons based on forecasting errors confirm that the Two-Moment CAPM with a time-varying systematic covariance risk is more efficient than the simpler time invariant model for forecasting returns on assets. In these literatures a number of different methods have emerged for modelling and forecasting time-varying systematic covariance risk in the conditional Two-Moment CAPM. Previously in Chapter 4, we compared the ability of GARCH-type models and Kalman Filter based approaches to model and forecast time-varying systematic covariance (*beta*) risk. The results suggested that the Kalman Filter based approaches outperform the GARCH-type models. The most well-known of these, the Kalman Filter Mean Reverting (KFMR) model (e.g. Wells (1996), Faff et al. (2000) and Mergner and Bulla (2008)), is employed again in this chapter.

The main purpose of this chapter is to assess the appropriateness of the Linear Market Model (allowing for only systematic covariance), which is consistent with the Two-Moment CAPM. We simultaneously compare its performance against six possible extensions, and such a comparison is yet to be undertaken in the literature. The first two are new reformulated forms of Higher order DGPs as simple polynomial extensions of the Linear Market Model, namely the Quadratic Market Model (allowing for systematic covariance and systematic skewness) and the Cubic Market Model (allowing for systematic covariance, systematic skewness and systematic kurtosis). The third approach relaxes some of the assumptions underpinning polynomial models by using a Generalized Additive Model (GAM), which is yet to be applied to assess the superiority of Higher order DGPs in finance. The last three approaches are the time-varying versions of the Linear Market Model and polynomial extensions in the form of state space models via KFMR; namely, the time-varying Linear Market Model (allowing for only time-varying systematic covariance), the time-varying Quadratic Market Model (allowing for time-varying systematic covariance and time-varying systematic skewness) and the time-varying Cubic Market Model (allowing for time-varying systematic covariance, time-varying systematic skewness and time-varying systematic kurtosis). The models are fitted by maximum likelihood, though in the case of the polynomial models a least squares estimation approach is equivalent. The aim of this chapter is to compare the performance of these models. In addition, the GAM is yet to be applied in the CAPM research in finance as well, so that is new. The Cubic Market Model is the most popular extension to the Linear Market Model, so we investigate whether it is really the best. In addition, the time-varying Higher order DGPs are possible extensions to the time-varying Linear Market Model. Thus, we examine whether these extensions are necessary to improve the model fit to the data. The comparison is made by using weekly data, generated by 9 developed and 9 emerging markets during three different time periods: the entire period from July 2002 to July 2012, from July 2002 to before the October 2008 financial crisis, and from after the October 2008 financial crisis to July 2012, thereby allowing one to investigate the effect of the October 2008 financial crisis when modelling stock market returns. In all cases the Morgan Stanley Capital International (MSCI) World Index and the three-month US dollar London Interbank Offered Rate (LIBOR) interest rate are used as a proxy for the market portfolio and the risk-free rate, respectively. The models are assessed by overall measures of model fit using *AIC*, *BIC* and *Adjusted* R^2, residual diagnostics, and by graphical summary of the fitted models to the data.

The rest of this chapter is outlined as follows. Section 5.2 outlines the seven models to be compared. Section 5.3 presents a description of the data. Section 5.4 presents the empirical results obtained from the model comparison during the entire period from July 2002 to July 2012, while section 5.5 presents the empirical results obtained from the model comparison during two different periods: from July 2002 to before the October 2008 financial crisis and from after the October 2008 financial crisis to July 2012. Section 5.6 presents our conclusions.

5.2 Methodology

5.2.1 Higher DGPs

The Four-Moment CAPM extends the Two-Moment CAPM by incorporating the systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*) in the data. The model is described in section 2.3 where more theoretical details are provided. To assess the necessity for the Four-Moment CAPM, the Cubic Market Model is fitted to the data in the form of a polynomial extension to the Linear Market Model. In this form, the Cubic Market Model is a third order polynomial in excess market returns, and can be written in the form

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1i}(R_{mt} - R_{ft}) + \alpha_{2i}(R_{mt} - R_{ft})^2 + \alpha_{3i}(R_{mt} - R_{ft})^3 + \varepsilon_{it}, \quad (5.1)$$

where R_{it} and R_{mt} are the stock market returns in country i and MSCI World market returns at time t ($t = 1, \ldots, T$), respectively. R_{ft} is the risk-free rate at time t. ε_{it} are the residuals with $\varepsilon_{it} \sim N(0, \sigma_i^2)$, $E(\varepsilon_{it}\varepsilon_{kt})$=0, for $i \neq k$ and

$E(\varepsilon_{it}\varepsilon_{i\,t+j})$=0, for $j > 0$. Here, the regression intercept, κ_i. Of the regression slopes, α_{1i} accounts for the systematic covariance (*proxy for* β_{im}), α_{2i} accounts for the systematic skewness (*proxy for* γ_{im}), and α_{3i} accounts for the systematic kurtosis (*proxy for* δ_{im}) and the proofs have been provided in section 2.3.3. Note that the Four-Moment CAPM is only appropriate if the DGP is at least cubic; that is if α_{3i} is statistically significantly different from zero. If not, then there will be collinearity in the systematic risk measures (β_{im}, γ_{im} and δ_{im}). The Quadratic Market Model ($\alpha_{3i} = 0$ in (5.1)) and the Linear Market Model ($\alpha_{2i} = 0$ and $\alpha_{3i} = 0$ in (5.1)) are reduced forms of the Cubic Market Model, and the latter is also the benchmark market model in finance. We have already illustrated that the Linear Market Model and the Quadratic Market Model are consistent with their equivalent CAPMs in section 2.1.2 and section 2.3.3.

5.2.2 Generalized Additive Model

The generalized additive model (GAM), was generated by Hastie and Tibshirani (1990), and is employed as a comparator to the polynomial models given by (5.1) to see if the latters rigid parametric shapes are too restrictive. In accordance with equation (3.14), the GAM function can be expressed as

$$R_{it} - R_{ft} = \kappa_i + f_i(R_{mt} - R_{ft}) + \varepsilon_{it}. \tag{5.2}$$

Here, $\varepsilon_{it} \sim N(0, \sigma_i^2)$ with $E(\varepsilon_{it}\varepsilon_{kt})$=0, for $i \neq k$, and $E(\varepsilon_{it}\varepsilon_{i\,t+j})$=0, for $j > 0$, and $f_i(R_{mt} - R_{ft})$ is a smooth function of $R_{mt} - R_{ft}$. The parameter estimation procedure used in generalized additive models is briefly outlined in section 3.2.

5.2.3 Time-varying Higher DGPs

In the financial literature, there now exists widespread evidence of the instability of the systematic risk measures, systematic covariance (*beta*), systematic skewness (*co-skewness*) and systematic kurtosis (*co-kurtosis*) in the Four Moment CAPM which is consistent with the Cubic Market Model. For the purpose of assessing this instability and to estimate its time-varying systematic risk measures, the

Cubic Market Model (5.1) can be extended to allow α_{1i}, α_{2i} and α_{3i} to evolve over time. This can be achieved using the model defined below, where estimation is via the Kalman Filter Mean Reverting (KFMR) algorithm which was outlined in section 3.3. The model has a state space form and is modified to become an observation equation expressed as

$$R_{it} - R_{ft} = \kappa_i + \alpha_{1it}(R_{mt} - R_{ft}) + \alpha_{2it}(R_{mt} - R_{ft})^2 + \alpha_{3it}(R_{mt} - R_{ft})^3 + \varepsilon_{it}. \quad (5.3)$$

Here, $\varepsilon_{it} \sim N(0, H_i)$. In accordance with equation (3.48), the state equations can be expressed as

$$\alpha_{1it} = \bar{\alpha}_{1i} + \phi_{1i}(\alpha_{1i\,t-1} - \bar{\alpha}_{1i}) + w_{1it}, \qquad w_{1it} \sim N(0, Q_{1i}), \quad (5.4)$$

$$\alpha_{2it} = \bar{\alpha}_{2i} + \phi_{2i}(\alpha_{2i\,t-1} - \bar{\alpha}_{2i}) + w_{2it}, \qquad w_{2it} \sim N(0, Q_{2i}), \quad (5.5)$$

$$\alpha_{3it} = \bar{\alpha}_{3i} + \phi_{3i}(\alpha_{3i\,t-1} - \bar{\alpha}_{3i}) + w_{3it}, \qquad w_{3it} \sim N(0, Q_{3i}), \quad (5.6)$$

with priors

$$\alpha_{1i0} \sim N(\mu_{\alpha_{1i}}, \Sigma_{\alpha_{1i}}), \quad \alpha_{2i0} \sim N(\mu_{\alpha_{2i}}, \Sigma_{\alpha_{2i}}), \quad \alpha_{3i0} \sim N(\mu_{\alpha_{3i}}, \Sigma_{\alpha_{3i}}), \quad (5.7)$$

where the parameters of these distributions estimated from the data as part of estimation algorithm.

Here, the regression intercept, κ_i and of the regression slopes, α_{1it} accounts for the time-varying systematic covariance (*proxy for* β_{imt}), α_{2it} accounts for the time-varying systematic skewness (*proxy for* γ_{imt}), and α_{3it} accounts for the time-varying systematic kurtosis (*proxy for* δ_{imt}). These are the extensions of the results proved in section 2.3.3. Note that the conditional Four-Moment CAPM is only appropriate if the time-varying DGP is at least cubic; that is if α_{3it} is statistically significantly different from zero. If not, then there will be collinearity in the time-varying systematic risk measures (β_{imt}, γ_{imt} and δ_{imt}).

Note that the time-varying Quadratic Market Model ($\alpha_{3it} = 0$ in (5.3)) and the time-varying Linear Market Model ($\alpha_{2it} = 0$ and $\alpha_{3it} = 0$ in (5.3)) are reduced forms of the time-varying Cubic Market Model. The time-varying Quadratic

Market Model (consistent with the conditional Three-Moment CAPM) and the time-varying Linear Market Model (consistent with the conditional Two-Moment CAPM) which are the extensions of the results proved in sections 2.3.3 and 2.1.2, respectively.

5.3 Data Description

All data for this study are weekly returns from 17 July 2002 to 18 July 2012 across 18 global market indices maintained by the Morgan Stanley Capital International Incorporation (MSCI Inc.). The data were obtained from the Thomson Reuters Financial Datastream database provided by University of Glasgow, UK. The 18 global markets include 9 developed markets: France, Germany, Italy, Japan, Norway, Sweden, Switzerland, the United Kingdom (UK) and the United States of America (USA); and 9 emerging markets: Brazil, Chile, India, Korea, Malaysia, Mexico, Poland, Russia and South Africa. The two criteria that were used to select the countries were: (1) regional classification in the database; and (2) continuous listing from 17 July 2002 to 18 July 2012. Table 5.1 represents an overview of all 18 global markets and their regions. The market proxy for this study is the Morgan Stanley Capital International World Market Index (MSCI) capturing 1,606 constituents including the large and mid cap representation across 24 developed markets countries (see further details at http://www.msci.com/).

The one-week returns for all 18 global markets and the MSCI World market portfolio were obtained from the first difference in the logarithm of Wednesday's closing price index as follows

$$R_{it} = \log(P_{it}) - \log(P_{i\,t-1}), \tag{5.8}$$

for $t = 2, \ldots, T$ and $i = 0, 1, \ldots, 18$, where $i = 0$ refers to the MSCI World market (R_{mt}) and $1 \leq i \leq 9$ refers to the developed markets, $10 \leq i \leq 18$ refers to the emerging markets. P_{it} is Wednesday's closing price index in week t. All indices are expressed in US dollars. The three-month US dollar London Interbank Offered Rate (LIBOR) interest rate served as a proxy for the risk-free rate, being

Table 5.1: Country stock market and regional classification.

Market	Region
Developed	
France	Europe & Middle East
Germany	Europe & Middle East
Italy	Europe & Middle East
Japan	Pacific
Norway	Europe & Middle East
Sweden	Europe & Middle East
Switzerland	Europe & Middle East
UK	Europe & Middle East
USA	Americas
Emerging	
Brazil	Americas
Chile	Americas
India	Asia
Korea	Asia
Malaysia	Asia
Mexico	Americas
Poland	Europe, Middle East & Africa
Russia	Europe, Middle East & Africa
South Africa	Europe, Middle East & Africa

Notes: Further details about country stock market and regional classifications which are maintained by the Morgan Stanley Capital International Incorporation (MSCI Inc.) are available at http://www.msci.com/products/indices/.

not constant through time in this research. As the LIBOR yields ($LIBOR_t$) are in percentage per annum, they can be converted to weekly rates as follows (Mergner (2009))

$$R_{ft} = \left(1 + \frac{LIBOR_t}{100}\right)^{1/52} - 1. \tag{5.9}$$

Table 5.2 shows descriptive statistics for the returns on the MSCI World market portfolio and the 18 global markets. The table provides some key points. The mean return in the emerging markets (0.0024) is higher than that of the developed markets (0.0007), which means that the emerging markets outperform the developed markets in this time period. The range of mean weekly returns varies from -0.0007 for Italy to 0.0033 for Brazil, meaning that Italy has a financial loss while Brazil has a financial gain on an investment during this period.

The average unconditional volatility of the returns in the emerging markets

Table 5.2: Descriptive statistics of weekly returns.

Market	Mean	Std.Dev.[a]	Skewness	Kurtosis	JB^b	$LB(23)^c$
World	0.0007	0.0264	-0.91	7.62	535.90*	200.08*
Developed						
France	0.0003	0.0376	-0.62	5.69	190.31*	195.38*
Germany	0.0007	0.0394	-0.84	6.09	268.87*	175.86*
Italy	-0.0007	0.0385	-0.53	5.73	186.32*	314.50*
Japan	0.0002	0.0281	-0.39	5.59	159.38*	138.98*
Norway	0.0017	0.0480	-1.04	8.06	650.17*	407.92*
Sweden	0.0018	0.0429	-0.73	6.08	253.39*	205.03*
Switzerland	0.0011	0.0285	-0.50	5.27	134.08*	156.26*
UK	0.0006	0.0317	-0.63	5.96	224.26*	335.73*
USA	0.0008	0.0258	-0.77	8.73	767.03*	162.09*
Average	*0.0007*	*0.0356*	*-0.67*	*6.36*		
Emerging						
Brazil	0.0033	0.0526	-1.68	13.50	2640.05*	47.37*
Chile	0.0032	0.0343	-1.80	19.08	5904.80*	17.54
India	0.0028	0.0431	-0.26	5.56	148.12*	228.78*
Korea	0.0017	0.0469	-0.42	10.09	1108.05*	521.15*
Malaysia	0.0018	0.0243	-0.16	5.46	134.40*	152.18*
Mexico	0.0027	0.0394	-1.99	19.05	5943.54*	45.60*
Poland	0.0015	0.0498	-1.08	7.07	461.36*	277.23*
Russia	0.0018	0.0587	-1.44	15.55	3604.33*	215.49*
South Africa	0.0025	0.0422	-0.89	7.61	530.74*	219.64*
Average	*0.0024*	*0.0435*	*-1.08*	*11.44*		
risk-free rate	0.0004	0.0003	0.63	1.90	60.96*	10762.17*

Notes: The portfolio has 522 observations for the weekly returns for each of the 18 global markets. [a]Std.Dev. is the standard deviation. [b]JB is the Jarque-Bera test statistic for testing the normality. JB follows χ^2 with 2 degrees of freedom and the critical value at the 5% level is 5.99. [c]LB statistic follows χ^2 with 23 degrees of freedom so the critical value at the 5% level is 35.17. * means the appropriate null hypothesis is rejected at the 5% significance level.

(0.0435) seems to be slightly higher than that of the developed markets (0.0356), which may be due to the variables capturing financial and economic integration, such as country credit risk ratings and the relative size of markets, being different in emerging markets as suggested by Bekaert and Harvey (1997). Therefore, the emerging markets are more risky in terms of the risk measured by unconditional volatility than the developed markets, which offsets their increased mean returns. The highest unconditional volatility (*standard deviation*) is that of the Russian market (0.0587) while the lowest one is that of Malaysia (0.0243).

The return distributions of the global markets and MSCI World market portfolio exhibit negative skewness, while the risk-free rate returns exhibit positive skewness. Negative skewness means that there are frequent small increases and a few extreme drops in returns, while positive skewness means that there are frequent small drops and a few extreme increases in returns (see Algieri (2012)). The range of skewness varies between -1.99 for Mexico and -0.16 for Malaysia. The average skewness of the returns in the emerging markets (-1.08) is also more negative than those of the developed markets (-0.67). This suggests that the developed markets have fewer extreme losses than the emerging markets when applied to investment returns.

The return distributions of the global markets, MSCI World market portfolio and risk-free rate are leptokurtic, meaning that the market has fatter tails than the normal distribution (which has kurtosis=3) and more chances of extreme outcomes. The range of kurtosis varies between 5.27 (Switzerland) and 19.08 (Chile). Emerging markets (11.44) demonstrate higher kurtosis than developed markets (6.36). This suggests that the emerging markets have more chance of extreme either financial losses or gains than the developed markets when applied to investment returns. The normality of each global market, MSCI World market portfolio and risk-free rate is also rejected at the 5% significance level using the Jarque-Bera (JB) test (see details in section 3.5.2.1) which is likely to be due to the substantial skewness and kurtosis observed from Table 5.2. According to the Ljung-Box ($LB(23)$) test, the null hypothesis of no autocorrelation for the squared returns (proxies for volatilities) is rejected at the 5% significance level for 18 global markets except for Chile, the MSCI World market portfolio and the risk-free rate, meaning that there exists a statistically significant autocorrelation for the squared returns which, in turn, provides strong evidence for the predictability of the volatility of 17 global markets, the MSCI World market portfolio and the risk-free rate (Christoffersen (2003)).

To sum up, the main features of these data in the developed and emerging markets are the positive mean (except for Italy), volatility, asymmetry (left-skew), and leptokurtosis (fat tails). The conclusions of Harvey's 1995 study match with the findings of this research, as he concluded that the most common features of stock market returns in emerging markets are high volatility, asymmetry, and leptokurtosis. This justifies giving consideration to higher moments, such as third and fourth moments, rather than just the Two-Moment CAPM for both developed and emerging markets.

Figure 5.1 displays time series plot of returns on the MSCI World market and 2 developed markets (UK and USA) and 2 emerging markets (Brazil and Russia), respectively. To save space the remaining markets are displayed in a smaller size in Figures 5.2 and 5.3. The figure provides some key points. In the whole time period, there is greater volatility in the emerging markets than in the developed markets and MSCI World market. Extreme events appear in October 2008, which is because in those weeks the financial crisis was starting to spread all over the world and October 6-10 turned out to be the worst week for stock markets in 75 years. It can be clearly seen that the emerging markets (Brazil and Russia) display even more volatility and extreme values in those weeks than the developed markets (UK and USA) and the MSCI World market, meaning that the emerging markets might be becoming increasingly integrated with this crisis than the developed markets. These extreme values can affect the modelling of asset returns. Hence, we choose three different time periods: the entire period of July 2002-July 2012, from July 2002 to before October 2008 and from after October 2008 to July 2012 which are analysed separately to investigate the effect of these extreme values while modelling asset returns.

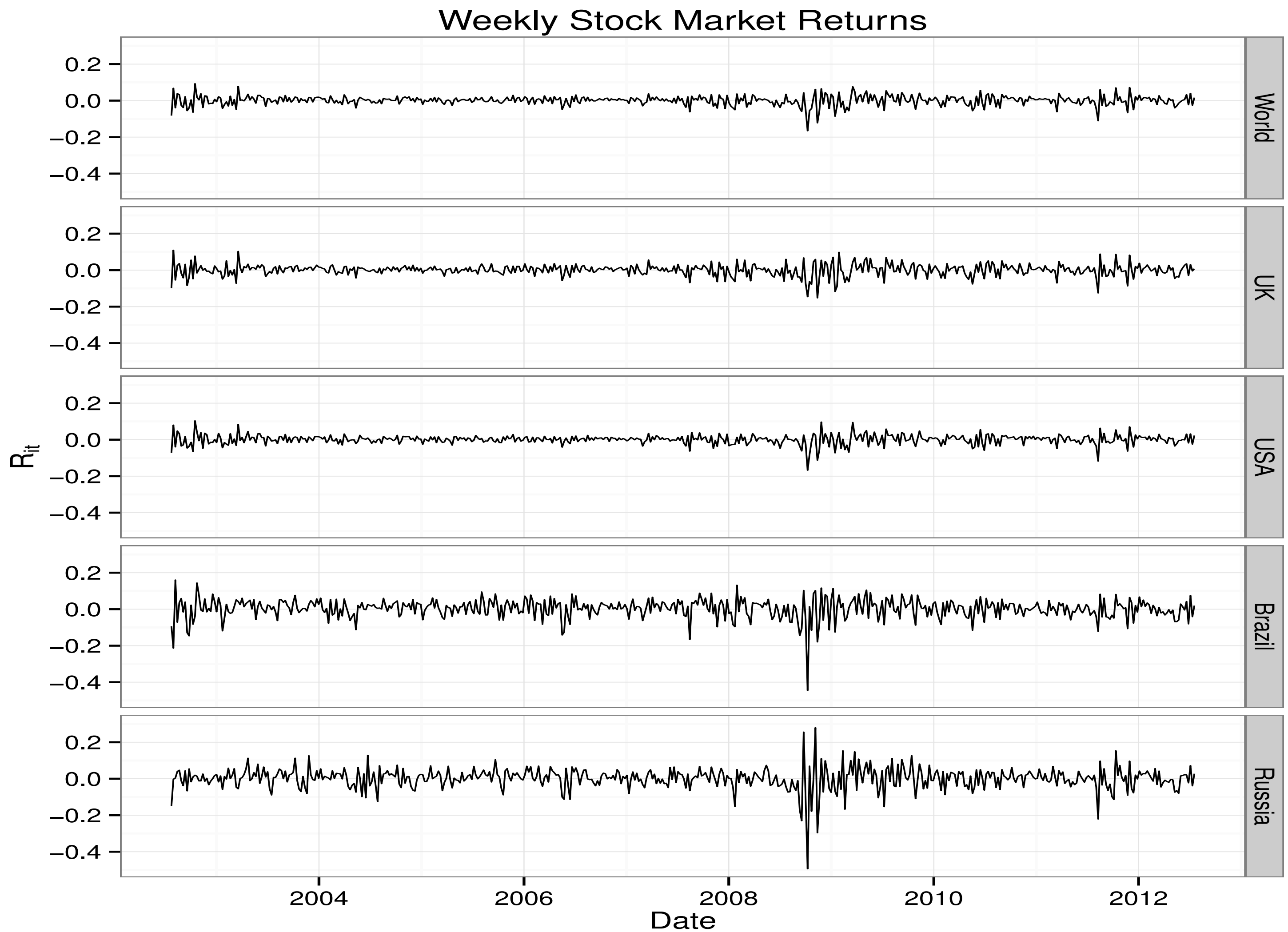


Figure 5.1: The time series plot of weekly returns on the MSCI World, 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets.

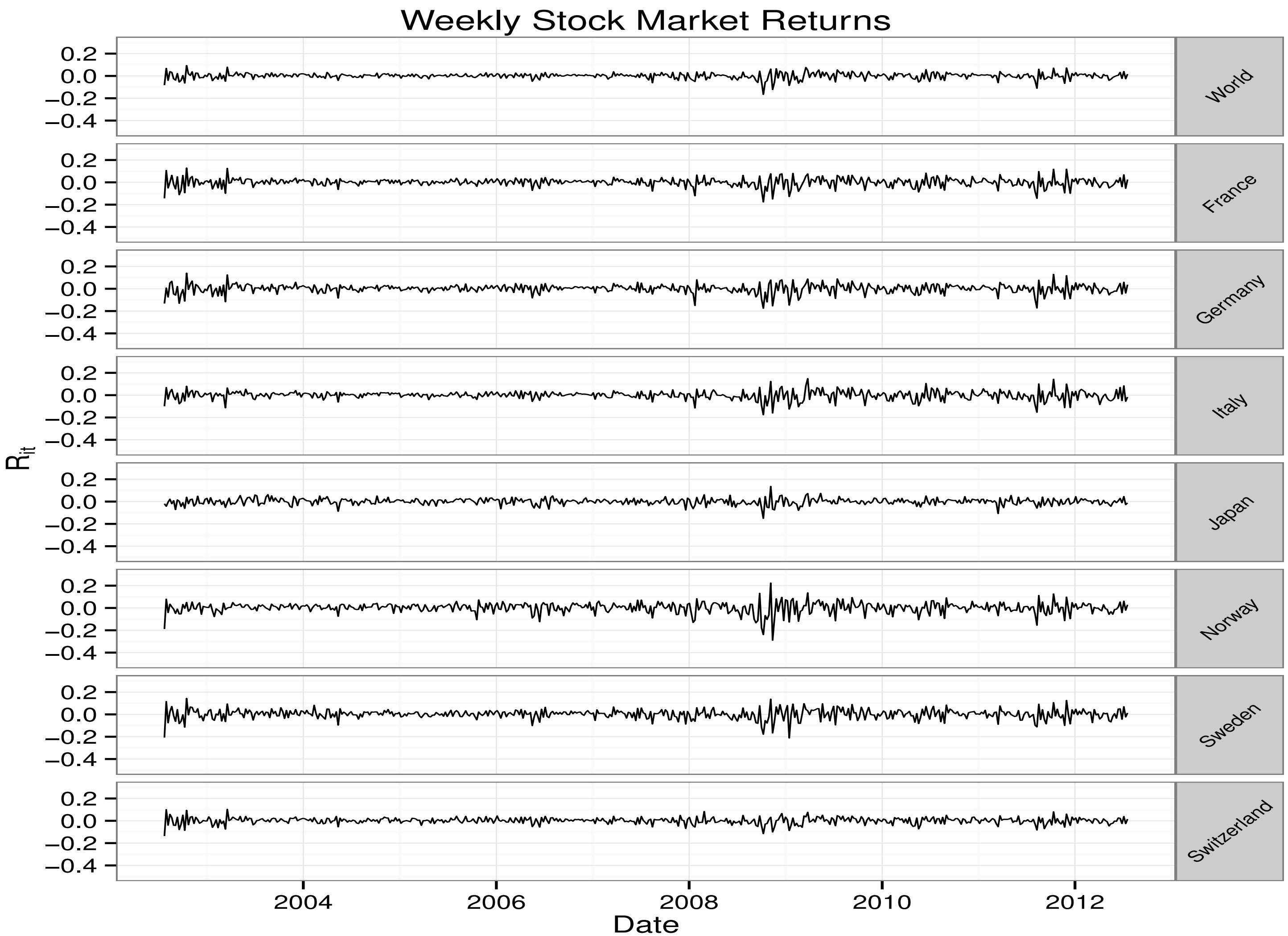


Figure 5.2: The time series plot of weekly returns on the MSCI World and 7 developed markets.

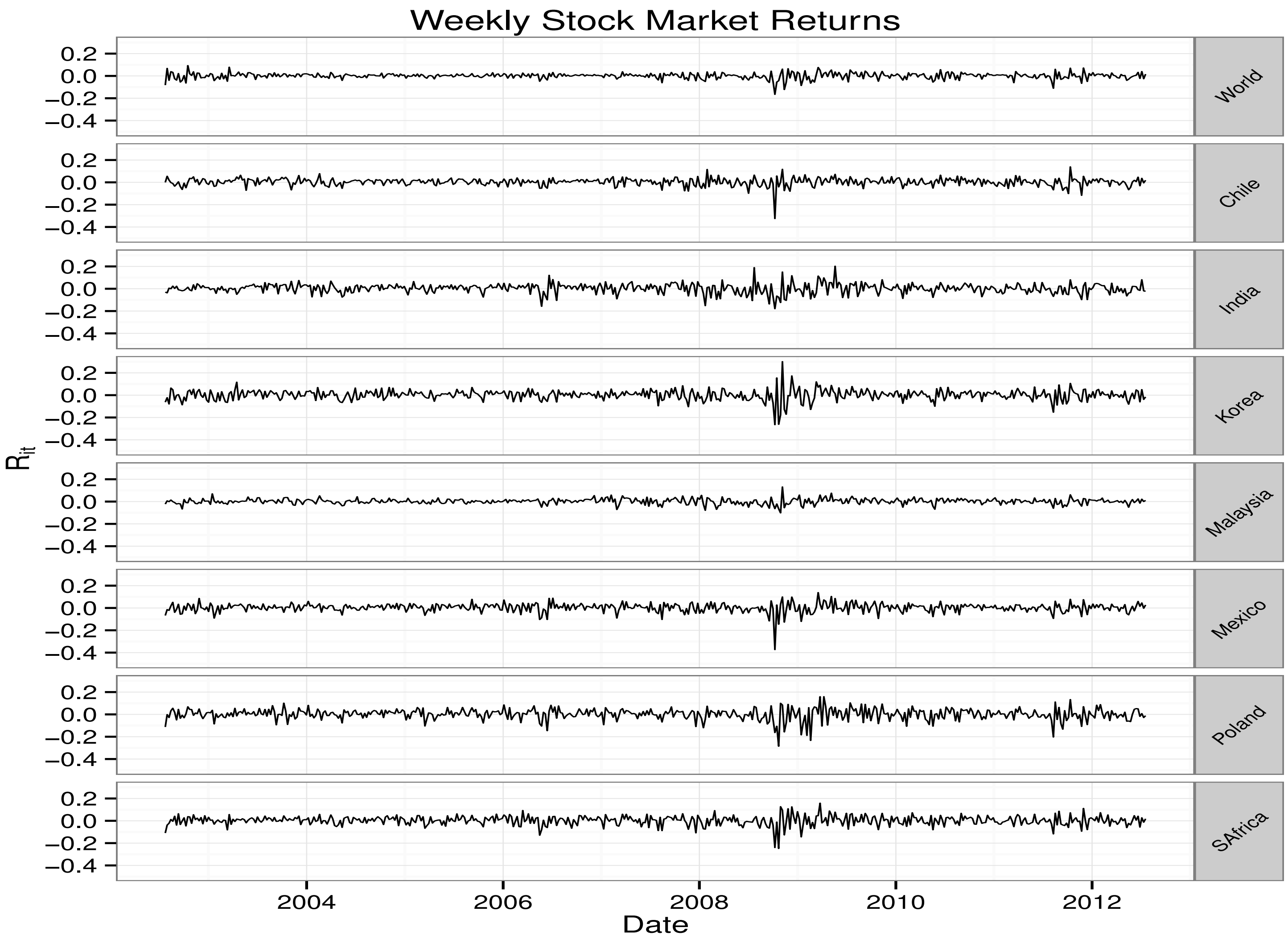


Figure 5.3: The time series plot of weekly returns on the MSCI World and 7 emerging markets.

5.4 Comparison of Models

5.4.1 Model Fit

In this section, a comparison is made between the Linear Market Model (LMM), the Quadratic Market Model (QMM), the Cubic Market Model (CMM), the generalized additive model (GAM), the time-varying Linear Market Model (TvLMM), the time-varying Quadratic Market Model (TvQMM) and the time-varying Cubic Market Model (TvCMM) for 18 global markets over the entire period from July 2002 to July 2012. The comparison is in terms of overall measures of model fit, using the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and *Adjusted* R^2 and residual diagnostics defined in section 3.5. A graphical comparison of the fit of each model to the data is also undertaken to determine the behaviour of the models. The model fit results are given in Tables 5.3, 5.4 and 5.5, which respectively display the AIC, BIC and *Adjusted* R^2.

Table 5.3: AIC values for all models & markets.

	AIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-2973.43	-2976.52	-2978.89	-2980.00	**-3019.84**	-3013.89	-3007.98
Germany	-2803.08	-2801.23	-2813.00	-2812.37	-2903.23	**-2904.75**	-2895.41
Italy	-2665.43	-2664.81	-2664.80	-2679.68	**-2702.24**	-2696.17	-2692.78
Japan	-2515.70	-2517.87	-2519.49	-2530.32	**-2554.98**	-2550.25	-2548.07
Norway	-2209.86	-2221.09	-2224.40	-2250.47	-2268.05	-2264.64	**-2271.88**
Sweden	-2582.33	-2581.50	-2588.09	-2594.19	**-2629.81**	-2624.56	-2618.88
Swit.	-2912.59	-2911.91	-2917.11	-2916.91	**-2939.59**	-2934.24	-2928.89
UK	-3044.62	-3042.94	-3045.70	-3054.95	**-3065.71**	-3059.75	-3054.02
USA	-3547.68	-3545.68	-3559.93	-3562.28	**-3606.95**	-3602.03	-3599.25
Emerging							
Brazil	-1954.09	-2001.36	-2001.16	-2028.72	-2116.05	**-2121.09**	-2115.92
Chile	-2332.24	-2347.09	-2354.47	-2405.01	-2418.98	**-2423.67**	-2418.97
India	-2038.91	-2037.51	-2040.67	-2049.69	**-2084.05**	-2081.06	-2078.68
Korea	-1993.95	-1995.51	-1993.58	-2001.04	**-2072.41**	-2066.42	-2060.86
Malaysia	-2609.18	-2607.40	-2618.40	-2621.46	**-2637.40**	-2632.02	-2629.74
Mexico	-2407.61	-2437.78	-2442.38	-2470.70	**-2518.91**	-2515.21	-2509.27
Poland	-1996.52	-1997.03	-2012.24	-2011.63	**-2059.04**	-2056.95	-2054.96
Russia	-1805.35	-1845.15	-1847.84	-1865.03	**-1916.32**	-1914.20	-1908.88
S.Africa	-2248.16	-2251.53	-2251.99	-2254.15	**-2281.66**	-2277.33	-2272.27

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest AIC.

Table 5.3 shows that the lowest AIC value comes from the time-varying Linear Market Model (TvLMM) via KFMR for all developed and emerging markets, with the exceptions being Germany, Brazil and Chile, where the time-varying Quadratic Market Model (TvQMM) via KFMR has the lowest AIC values, as well as Norway, where the time-varying Cubic Market Model (TvCMM) via KFMR has the lowest AIC value. In addition, the AIC value of the time-varying Quadratic Market Model is lower than that of the time-varying Cubic Market Model (TvCMM) for all 18 global markets except for Norway, suggesting that the model fit performance of the time-varying Cubic Market Model is worse than that of other time-varying DGPs. The time-varying Cubic Market Model outperforms the Higher order DGPs and the generalized additive model (GAM) for all developed and emerging markets, with the exception being the UK, where the GAM is better. In addition, the GAM outperforms the Higher order DGPs in general across the 18 global markets, with the exceptions being Switzerland and

Table 5.4: BIC values for all models & markets.

	BIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-2960.66	-2959.49	-2957.61	-2951.83	**-2998.55**	-2979.82	-2961.15
Germany	-2790.30	-2784.20	-2791.71	-2782.89	**-2881.94**	-2870.69	-2848.58
Italy	-2652.65	-2647.78	-2643.51	-2637.88	**-2680.96**	-2662.11	-2645.94
Japan	-2502.92	-2500.84	-2498.20	-2500.29	**-2533.69**	-2516.19	-2501.24
Norway	-2197.09	-2204.06	-2203.11	-2205.78	**-2246.76**	-2230.58	-2225.05
Sweden	-2569.56	-2564.47	-2566.80	-2552.05	**-2608.52**	-2590.49	-2572.05
Swit.	-2899.82	-2894.88	-2895.82	-2890.55	**-2918.30**	-2900.18	-2882.05
UK	-3031.85	-3025.91	-3024.41	-3013.69	**-3044.42**	-3025.69	-3007.18
USA	-3534.90	-3528.65	-3538.65	-3527.01	**-3585.66**	-3567.97	-3552.42
Emerging							
Brazil	-1941.31	-1984.33	-1979.87	-1988.11	**-2094.77**	-2087.03	-2069.09
Chile	-2319.47	-2330.06	-2333.18	-2362.57	**-2397.69**	-2389.61	-2372.14
India	-2026.14	-2020.48	-2019.39	-2013.33	**-2064.76**	-2047.00	-2031.85
Korea	-1981.18	-1978.48	-1972.29	-1966.85	**-2051.12**	-2032.36	-2014.03
Malaysia	-2596.41	-2590.37	-2597.11	-2578.70	**-2616.11**	-2597.96	-2582.91
Mexico	-2394.83	-2420.75	-2421.09	-2431.38	**-2497.63**	-2481.15	-2462.44
Poland	-1983.75	-1980.00	-1990.95	-1972.88	**-2037.75**	-2022.89	-2008.13
Russia	-1792.58	-1828.12	-1826.55	-1821.01	**-1895.03**	-1880.14	-1862.05
S.Africa	-2235.39	-2234.50	-2230.70	-2215.23	**-2260.37**	-2243.27	-2225.44

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest BIC.

Poland, where the Cubic Market Model (CMM) is better. Also, the Linear Market Model (LMM) outperforms the Quadratic Market Model (QMM) in Germany, Italy, Sweden, Switzerland, the UK, the USA, India and Malaysia.

Table 5.4 shows that, based on the lowest *BIC*, the time-varying Linear Market Model (TvLMM) via KFMR is the most appropriate model for all developed and emerging markets. Among the time-varying DGPs, the time-varying Quadratic Market Model (TvQMM) has lower *BIC* values than the time-varying Cubic Market Model (TvCMM) for all 18 global markets, suggesting that the time-varying Quadratic Market Model is preferable to the time-varying Cubic Market Model. It can be seen that the Linear Market Model outperforms the time-varying Quadratic Market Model for the UK and the time-varying Cubic Market Model for 4 developed markets, Italy, Japan, Switzerland and the UK as well as 2 emerging markets, Malaysia and South Africa. The Linear Market Model (LMM) also outperforms the other Higher order DGPs and the GAM for 6 developed markets, France, Italy, Japan, Sweden, Switzerland and the UK, as well as 3 emerging markets, India, Korea and South Africa. This is because the *BIC* has a larger penalty than the *AIC*, and thus chooses simpler models. Also, the GAM outperforms the Higher order DGPs for only 1 developed market, Norway, as well as 3 emerging markets, Brazil, Chile, and Mexico. These results show that the time-varying Linear Market Model outperforms all other models, even in terms of *BIC* which prefers simplier models than *AIC*. This suggests that it represents a substantial improvement in model fit compared with all other models. In contrast, the remaining models sometimes do not fit the data as well as the Linear Market Model as measured by *BIC*, suggesting that their additional complexity does not greatly improve model fit.

The results for *Adjusted* R^2 are given in Table 5.5. Again, this shows that the time-varying Linear Market Model (TvLMM) provides a better performance than the time-varying Higher order DGPs for all 18 global markets, with the exception of France, Italy, Norway, the USA, India and Poland, where the time-varying Quadratic Market Model has an equal or a better performance to that of the time-varying Linear Market Model and of Germany, where the time-varying Cubic Market Model has an equal performance to that of the time-varying Linear

Table 5.5: *Adjusted* R^2 values for all models & markets.

	Adjusted R^2						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	0.862	0.863	0.864	0.865	**0.921**	**0.921**	0.920
Germany	0.826	0.825	0.830	0.830	**0.928**	0.909	**0.928**
Italy	0.763	0.763	0.764	0.772	**0.868**	**0.868**	0.866
Japan	0.407	0.411	0.414	0.428	**0.661**	0.658	0.655
Norway	0.634	0.642	0.645	0.666	0.772	**0.780**	0.746
Sweden	0.775	0.775	0.779	0.783	**0.869**	0.863	0.868
Swit.	0.730	0.731	0.734	0.734	**0.814**	0.812	0.810
UK	0.831	0.831	0.832	0.836	**0.883**	0.881	0.882
USA	0.903	0.903	0.905	0.906	**0.954**	**0.954**	0.953
Emerging							
Brazil	0.502	0.546	0.547	0.574	**0.822**	0.810	0.811
Chile	0.435	0.452	0.461	0.515	**0.698**	0.653	0.649
India	0.370	0.369	0.374	0.389	0.525	**0.542**	0.522
Korea	0.422	0.424	0.423	0.435	**0.745**	0.744	0.743
Malaysia	0.334	0.333	0.349	0.358	**0.607**	0.606	0.595
Mexico	0.628	0.650	0.654	0.675	**0.845**	0.843	0.843
Poland	0.489	0.491	0.506	0.510	0.753	**0.754**	0.750
Russia	0.469	0.509	0.512	0.533	**0.740**	0.733	0.734
S.Africa	0.560	0.564	0.565	0.570	**0.752**	0.749	0.747

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the highest *Adjusted* R^2.

Market Model. Nevertheless, the improvements in model fit to the time-varying Higher order DGPs are not substantial. For example, the time-varying Linear Market Model improves on the time-varying Quadratic Market Model in terms of *Adjusted* R^2 by on average 0.3% for the developed markets and 0.6% for the emerging markets. The time-varying Linear Market Model also improves on the time-varying Cubic Market Model in terms of *Adjusted* R^2 by on average 0.5% for the developed markets and 1% for the emerging markets. In addition, the time-varying Quadratic Market Model outperforms the time-varying Cubic Market Model for 12 out of 18 global markets, and it improves on the time-varying Cubic Market Model in terms of *Adjusted* R^2 by on average 0.2% for the developed markets and 0.4% for the emerging markets. These results confirm what is seen in Tables 5.3 and 5.4, namely, that the additional complexity of the time-varying Higher order DGPs does not improve model fit compared to the time-varying Linear Market Model.

The time-varying Linear Market Model via KFMR provides a much better performance than the Higher order DGPs and the GAM. The time-varying Linear Market Model improves on the Linear Market Model in terms of *Adjusted* R^2 by on average 10.4% for the developed markets and 25.3% for the emerging markets. This suggests that emerging markets are more unstable than the developed markets. In addition, the GAM outperforms the Higher order DGPs in general across the 18 global markets, but the improvements in model fit are not substantial. For example, the GAM improves on the Linear Market Model in terms of *Adjusted* R^2 by on average only 1% for the developed markets and 3.9% for the emerging markets. The average increase in *Adjusted* R^2 from the Linear Market Model to the Cubic Market Model is 0.4% for the developed markets and 2% for the emerging markets. Indeed, the Linear Market Model outperforms the Quadratic Market Model in Germany, India and Malaysia in terms of *Adjusted* R^2. These results confirm what is seen in Tables 5.3 and 5.4, namely that the time-varying Linear Market Model offers a substantial improvement in model fit to the Linear Market Model, but that the other non-linear DGPs do not. In addition, this improvement in model fit is most substantial for the emerging markets, suggesting they are less able to be charactered by a straight line.

5.4.2 Residual Diagnostics

Diagnostic test statistics are provided in Tables 5.6, 5.7 and 5.8 outlined in section 3.5.2.1. According to the Jarque-Bera (JB) test, the residuals are not normally distributed at the 5% significance level for all markets and models, implying that all models are poor in terms of non-normal errors. The time-varying DGPs via KFMR have not solved this. According to the Ljung-Box ($LB(23)$) test, the null hypothesis of no autocorrelation can be rejected at the 5% significance level for most markets and models. Although autocorrelation remains in most data sets after fitting most models, it is vastly reduced using the time-varying DGPs via KFMR model. Average decrease in the $LB(23)$ test statistics for the time-varying Linear Market Model (TvLMM) via KFMR model compared with the Cubic Market Model (CMM) are 38.6% for the developed markets and 22.6% for

Table 5.6: Normality test statistics for all models & markets.

	JB						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	102.07*	118.72*	105.97*	113.38*	261.89*	261.37*	259.85*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Germany	294.06*	304.79*	207.74*	222.41*	100.54*	32.87*	90.37*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Italy	45.93*	40.50*	46.51*	32.64*	109.24*	108.27*	102.12*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Japan	67.18*	68.79*	78.77*	106.71*	87.98*	86.15*	87.56*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Norway	133.22*	116.82*	139.58*	72.24*	78.68*	83.59*	75.18*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Sweden	114.67*	130.52*	112.57*	64.10*	93.02*	78.38*	86.64*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Swit.	33.59*	40.53*	33.67*	32.22*	33.14*	31.11*	30.39*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
UK	141.63*	138.33*	137.64*	151.64*	312.38*	296.67*	297.09*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
USA	88.34*	88.13*	107.05*	97.22*	44.43*	43.31*	40.82*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Emerging							
Brazil	4383.99*	2005.36*	1821.74*	1948.91*	66.93*	56.38*	53.76*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Chile	817.98*	124.25*	62.88*	59.35*	55.92*	34.08*	32.88*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
India	132.74*	134.65*	131.79*	115.64*	105.81*	133.65*	134.71*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Korea	1900.93*	2163.24*	2192.52*	2337.99*	488.12*	489.66*	493.80*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Malaysia	229.48*	230.47*	293.06*	268.82*	304.55*	308.31*	317.86*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Mexico	1054.32*	75.38*	52.96*	44.84*	26.92*	26.71*	27.10*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Poland	513.14*	498.66*	413.96*	455.74*	57.47*	56.48*	50.99*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
Russia	403.55*	297.95*	274.03*	169.71*	247.08*	218.84*	221.52*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*	*(0.000)*
S.Africa	616.13*	598.02*	535.16*	612.09*	14.24*	13.69*	13.23*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.001)*	*(0.001)*	*(0.001)*

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. *JB* is the Jarque-Bera test statistic for the null hypothesis of normally distributed standardised residuals. *JB* follows χ^2 with 2 degrees of freedom so the critical value at the 5% level is 5.99. * means the null hypothesis of normality is rejected at the 5% significance level.

Table 5.7: Autocorrelation test statistics for all models & markets.

	$LB(23)$						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	35.26*	37.61*	38.79*	37.32*	24.49	24.41	24.34*
	(0.049)	*(0.028)*	*(0.021)*	*(0.030)*	*(0.178)*	*(0.081)*	*(0.028)*
Germany	29.69	29.36	27.61	27.35	17.50	15.04	17.55
	(0.159)	*(0.169)*	*(0.231)*	*(0.241)*	*(0.556)*	*(0.522)*	*(0.175)*
Italy	15.46	15.85	17.02	17.18	9.80	9.89	10.50
	(0.877)	*(0.862)*	*(0.808)*	*(0.800)*	*(0.958)*	*(0.872)*	*(0.653)*
Japan	28.46	32.16	33.17	32.43	16.94	17.73	18.09
	(0.199)	*(0.097)*	*(0.078)*	*(0.092)*	*(0.594)*	*(0.340)*	*(0.154)*
Norway	60.44*	67.82*	68.65*	59.80*	32.88*	31.51*	32.88*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.025)*	*(0.012)*	*(0.002)*
Sweden	72.31*	71.65*	64.02*	60.01*	34.37*	33.77*	33.29*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.017)*	*(0.006)*	*(0.002)*
Swit.	32.05	31.83	36.06*	33.73	29.93	29.46*	29.32*
	(0.099)	*(0.104)*	*(0.041)*	*(0.069)*	*(0.053)*	*(0.021)*	*(0.006)*
UK	58.57*	58.68*	53.62*	52.40*	41.34*	41.41*	41.15*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.002)*	*(0.000)*	*(0.000)*
USA	55.58*	55.63*	56.34*	53.43*	31.17*	30.51*	30.63*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.039)*	*(0.016)*	*(0.004)*
Emerging							
Brazil	65.37*	43.18*	42.71*	34.16	19.73	21.41	22.23
	(0.000)	*(0.007)*	*(0.007)*	*(0.063)*	*(0.411)*	*(0.163)*	*(0.052)*
Chile	23.98	18.79	16.81	21.96	21.15	18.93	18.97
	(0.405)	*(0.713)*	*(0.818)*	*(0.522)*	*(0.329)*	*(0.272)*	*(0.124)*
India	36.82*	37.28*	40.88*	40.74*	34.11*	33.42*	36.87*
	(0.034)	*(0.030)*	*(0.012)*	*(0.013)*	*(0.018)*	*(0.006)*	*(0.000)*
Korea	108.26*	117.69*	116.85*	119.68*	18.70	18.78	18.80
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.476)*	*(0.280)*	*(0.129)*
Malaysia	21.05	21.61	19.82	20.03	24.29	24.37	23.71*
	(0.578)	*(0.544)*	*(0.653)*	*(0.640)*	*(0.185)*	*(0.082)*	*(0.034)*
Mexico	56.09*	42.84*	39.69*	36.35*	29.18	29.90*	29.84*
	(0.000)	*(0.007)*	*(0.017)*	*(0.038)*	*(0.063)*	*(0.019)*	*(0.005)*
Poland	57.00*	56.81*	55.36*	54.11*	41.28*	38.77*	38.13*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.002)*	*(0.001)*	*(0.000)*
Russia	32.29	32.75	30.65	26.09	25.09	25.03	24.97*
	(0.094)	*(0.086)*	*(0.132)*	*(0.297)*	*(0.158)*	*(0.069)*	*(0.023)*
S.Africa	47.84*	42.26*	43.60*	41.16*	31.66*	31.10*	30.95*
	(0.002)	*(0.008)*	*(0.006)*	*(0.011)*	*(0.034)*	*(0.013)*	*(0.003)*

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. *LB(23)* is the Ljung-Box test statistic for the null hypothesis of no autocorrelation in the standardised residuals up to order $\sqrt{522} \approx 23$. In the DGPs and GAM, LB statistic follows χ^2 with 23 degrees of freedom so the critical value at the 5% level is 35.17 (Mergner (2009)). In the time-varying DGPs via KFMR, LB statistic follows χ^2 with 23-$(m-1)$ degrees of freedom where m is the total number of estimated parameters (Harvey (1989)). * means the null hypothesis of no autocorrelation is rejected at the 5% significance level.

Table 5.8: Heteroskedasticity test statistics for all models & markets.

	Het(174)						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	0.87	0.84	0.84	0.87	0.72	0.72	0.72
	(0.820)	*(0.875)*	*(0.875)*	*(0.820)*	*(0.984)*	*(0.984)*	*(0.984)*
Germany	0.80	0.80	0.83	0.85	0.69	0.79	0.70
	(0.930)	*(0.930)*	*(0.890)*	*(0.858)*	*(0.993)*	*(0.939)*	*(0.990)*
Italy	1.85*	1.75*	1.78*	1.87*	1.53*	1.52*	1.54*
	(0.000)	*(0.000)*	*(0.000)*	*(0.000)*	*(0.003)*	*(0.003)*	*(0.002)*
Japan	0.81	0.82	0.85	0.88	0.76	0.76	0.78
	(0.917)	*(0.904)*	*(0.858)*	*(0.800)*	*(0.964)*	*(0.964)*	*(0.948)*
Norway	0.50	0.58	0.60	0.61	0.59	0.57	0.67
	(0.999)	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.996)*
Sweden	1.00	0.99	0.98	1.02	0.84	0.85	0.84
	(0.500)	*(0.526)*	*(0.552)*	*(0.448)*	*(0.874)*	*(0.858)*	*(0.874)*
Swit.	0.78	0.79	0.78	0.79	0.73	0.73	0.73
	(0.948)	*(0.939)*	*(0.948)*	*(0.939)*	*(0.980)*	*(0.980)*	*(0.980)*
UK	0.59	0.58	0.58	0.61	0.57	0.57	0.57
	(0.999)	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*
USA	0.97	0.97	1.03	1.00	0.77	0.78	0.78
	(0.579)	*(0.579)*	*(0.423)*	*(0.500)*	*(0.955)*	*(0.949)*	*(0.949)*
Emerging							
Brazil	0.23	0.31	0.33	0.30	0.46	0.47	0.47
	(0.999)	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*
Chile	0.88	1.02	0.95	1.01	0.72	0.79	0.79
	(0.800)	*(0.448)*	*(0.632)*	*(0.474)*	*(0.985)*	*(0.939)*	*(0.939)*
India	1.04	1.07	1.08	1.06	1.06	1.03	1.05
	(0.398)	*(0.328)*	*(0.306)*	*(0.350)*	*(0.350)*	*(0.423)*	*(0.374)*
Korea	0.76	0.79	0.79	0.80	0.63	0.63	0.64
	(0.964)	*(0.939)*	*(0.939)*	*(0.929)*	*(0.999)*	*(0.999)*	*(0.998)*
Malaysia	0.76	0.77	0.82	0.82	0.63	0.63	0.64
	(0.964)	*(0.957)*	*(0.904)*	*(0.904)*	*(0.999)*	*(0.999)*	*(0.998)*
Mexico	0.63	0.76	0.71	0.72	0.60	0.60	0.61
	(0.998)	*(0.964)*	*(0.987)*	*(0.985)*	*(0.999)*	*(0.999)*	*(0.999)*
Poland	0.91	0.95	0.98	0.96	0.77	0.78	0.79
	(0.732)	*(0.632)*	*(0.552)*	*(0.606)*	*(0.957)*	*(0.949)*	*(0.939)*
Russia	0.39	0.49	0.47	0.47	0.38	0.39	0.39
	(0.999)	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*	*(0.999)*
S.Africa	0.74	0.80	0.83	0.81	0.74	0.75	0.75
	(0.976)	*(0.929)*	*(0.890)*	*(0.917)*	*(0.976)*	*(0.970)*	*(0.970)*

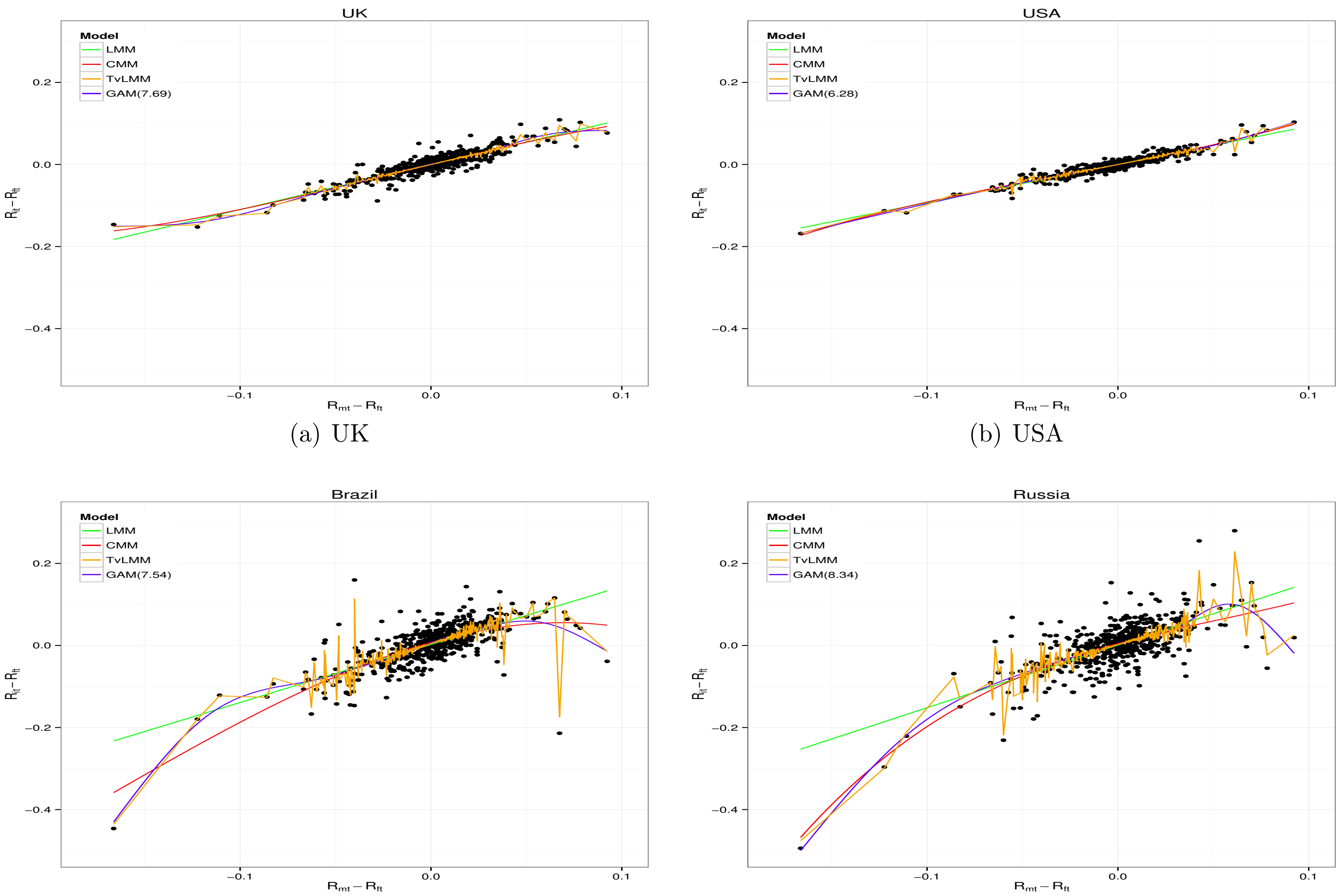
Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. *Het(174)* is the test statistic for the null hypothesis of no heteroskedasticity in the standardised residuals up to order $522/3 = 174$. *Het(174)* statistic follows $F_{(174,174)}$ distribution so the critical value at the 5% level is 1.28. * means the null hypothesis of no heteroskedasticity is rejected at the 5% significance level.

the emerging markets. According to the H ($Het(174)$) test, the null hypothesis of no heteroskedasticity cannot be rejected for all 18 global markets except for Italy and models at the 5% significance level, implying that all models are adequate in terms of no heteroskedasticity.

5.4.3 Graphical Summary

Figure 5.4 presents the scatter plots of the relationship between each stock market excess return and the MSCI World market excess return for 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets, respectively. To save space the remaining markets are displayed in a smaller size in Figures 5.5 and 5.6. These plots include the fitted models for the Linear Market Model (LMM), the Cubic Market Model (CMM), the GAM function and the time-varying Linear Market Model (TvLMM) via KFMR. The figure provides some key points. The fitted curve for the Quadratic Market Model (QMM) is similar to that of the Cubic Market Model (CMM). Also, the fitted curves for the time-varying Quadratic Market Model (TvQMM) and the time-varying Cubic Market Model (TvCMM) are similar to the time-varying Linear Market Model (TvLMM). These are not shown for ease of presentation.

It can be seen that the weekly excess market returns of both developed and emerging markets are positively correlated with the MSCI World market weekly excess returns, with a correlation coefficient ranging from 0.58 (Malaysia) to 0.95 (USA). The time-varying Linear Market Model (TvLMM) provides a much closer fit to the data than the others due to the time-varying relationship estimated between $R_{it} - R_{ft}$ and $R_{mt} - R_{ft}$. Such short-term volatility in this relationship cannot be captured by the globally smooth polynomial or generalized additive models. The figure also suggests that the Linear Market Model (LMM) can be appropriate for capturing the risk-return relationship without extreme values in the data sample, as the majority of the UK and USA data exhibit a close to linear relationship. Note that the non-linear models exhibit estimated relationships that are close to that of the linear one. This is not true for the emerging markets however. Finally, the curvature observed from the GAM fit appears to be driven



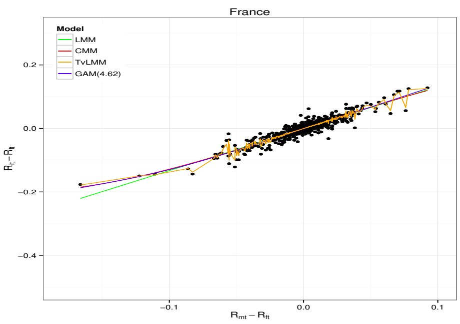
(a) UK

(b) USA

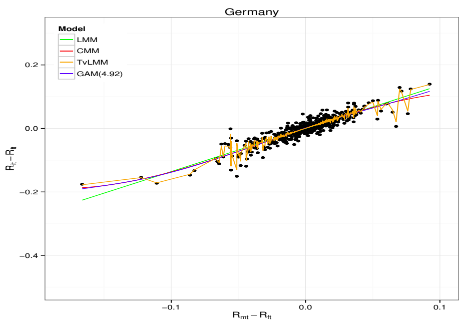
(c) Brazil

(d) Russia

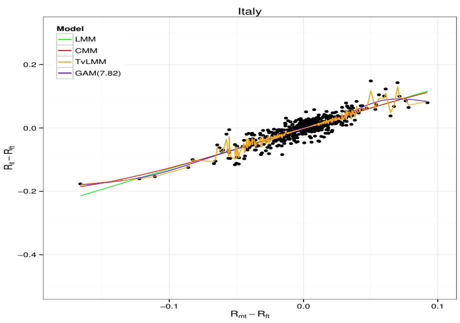
Figure 5.4: The scatter plots of 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets weekly excess returns.



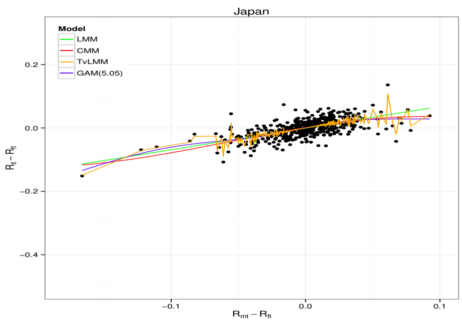
(a) France



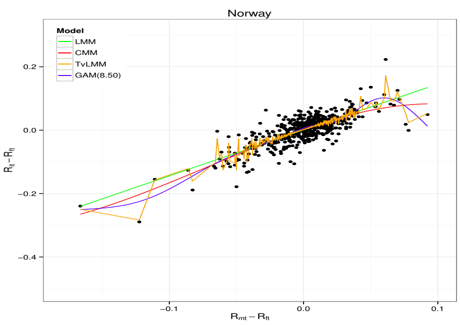
(b) Germany



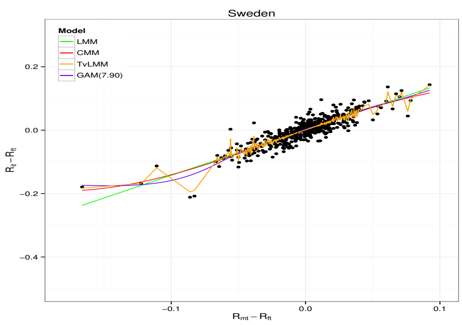
(c) Italy



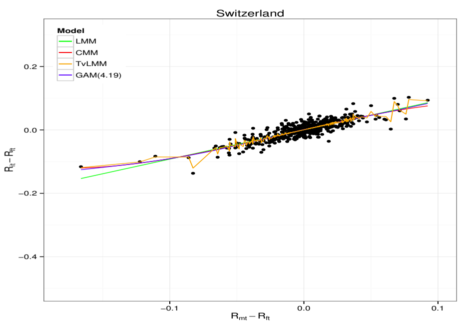
(d) Japan



(e) Norway

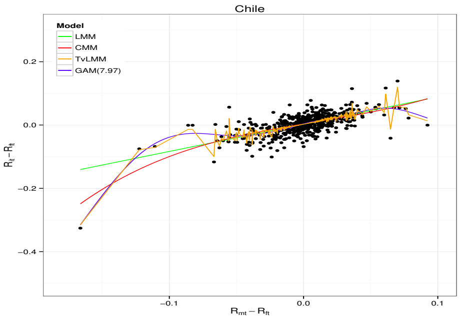


(f) Sweden

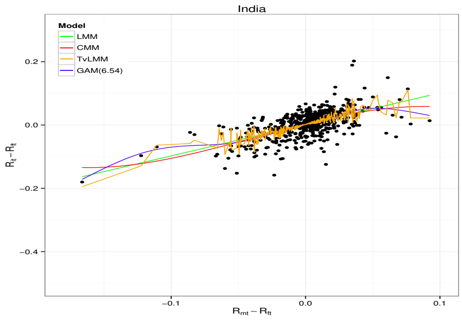


(g) Switzerland

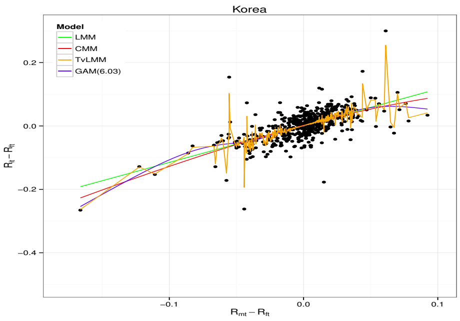
Figure 5.5: The scatter plots of 7 developed markets weekly excess returns.



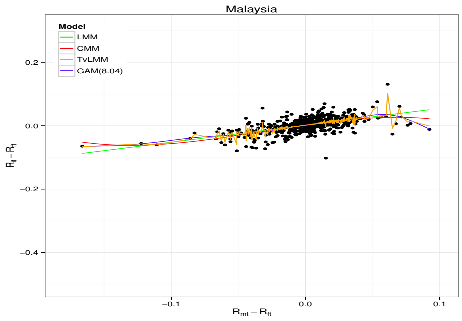
(a) Chile



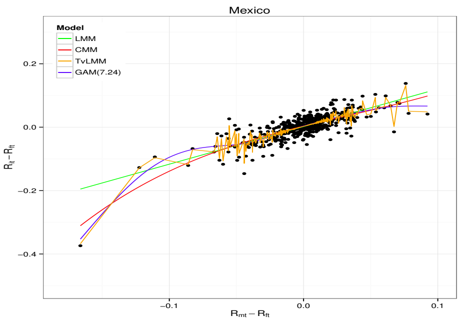
(b) India



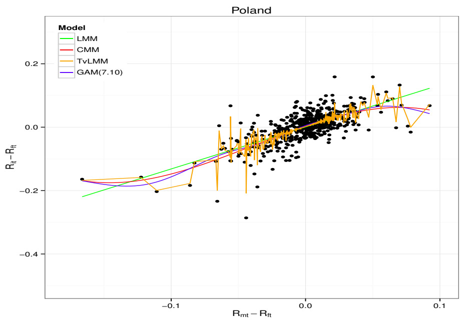
(c) Korea



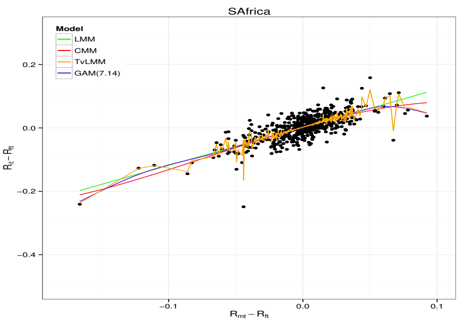
(d) Malaysia



(e) Mexico



(f) Poland



(g) South Africa

Figure 5.6: The scatter plots of 7 emerging markets weekly excess returns.

largely by a small number of outlying values, as it appears roughly linear for the vast majority of the data.

To sum up, the time-varying Linear Market Model (TvLMM) via KFMR with time-varying systematic covariance (*beta*) risk appears to be the most appropriate model for developed and emerging markets in terms of overall measures of model fit and the graphical summary of the fitted models to the data.

5.4.4 Time-varying Linear Market Model

As the time-varying Linear Market Model (TvLMM) exhibits the best fit to the data, we examine it in greater detail here. Table 5.9 presents hyperparameter estimates from the time-varying Linear Market Model via KFMR, which has been applied using equations (5.3) and (5.4).

The parameter ϕ_{1i}, which is crucial to clarify the temporal autocorrelation in $\{\alpha_{1it}\}$ within the Kalman Filter based approaches, should lie in the range 0 to 1 for a stationary series (see details in section 3.3). Here, estimated $\hat{\phi}_{1i}$ is close to 1 for India, so that the time-varying systematic covariance (*beta*) series of the KFMR becomes similar to the KFRW. On the other hand, the time-varying systematic covariance series of the KFMR becomes similar to the KFRC when $\hat{\phi}_{1i}$ is close to 0 (for Germany, Japan, Sweden and USA in the developed markets and Chile, Korea, Malaysia, and Mexico in the emerging markets).

Across all 18 global markets, the estimated values of $\hat{Q}_{1i}$ are higher than those of the $\hat{H}_i$, meaning that the state variance captures the volatility of the stock market excess returns more than the observation variance. Also, both the estimated $\hat{H}_i$ and $\hat{Q}_{1i}$ values for the emerging markets are generally higher than those of the developed markets, meaning that the emerging markets are more volatile than the developed markets.

Table 5.10 presents state parameter estimates from the time-varying Linear Market Model via KFMR (equations (5.3) and (5.4)) in both developed and emerging markets.

The regression intercept, $\hat{\kappa}_i$, of all 18 global markets is close to zero. This is an expected result for $\hat{\kappa}_i$ in the Two-Moment CAPM because the risk-free rate

Table 5.9: Time-varying Linear Market Model hyperparameter estimates (standard errors) via KFMR.

Market	$\hat{Q}_{1i} \times 100$	$\hat{H}_i \times 100$	$\hat{\phi}_{1i}$
Developed			
France	8.081	0.014	0.223
	(2.106)	*(0.001)*	*(0.053)*
Germany	16.920	0.015	0.029
	(2.917)	*(0.001)*	*(0.011)*
Italy	17.249	0.024	0.119
	(4.223)	*(0.002)*	*(0.032)*
Japan	21.197	0.033	0.057
	(5.111)	*(0.003)*	*(0.028)*
Norway	20.077	0.061	0.575
	(12.042)	*(0.005)*	*(0.339)*
Sweden	16.739	0.029	0.000
	(4.110)	*(0.002)*	*(0.000)*
Switzerland	5.873	0.017	0.278
	(2.167)	*(0.001)*	*(0.098)*
UK	4.433	0.014	0.418
	(1.810)	*(0.001)*	*(0.135)*
USA	3.914	0.004	0.020
	(0.811)	*(0.000)*	*(0.009)*
Emerging			
Brazil	70.647	0.066	0.436
	(14.591)	*(0.006)*	*(0.081)*
Chile	26.008	0.043	0.000
	(5.871)	*(0.003)*	*(0.000)*
India	5.204	0.095	0.927
	(3.695)	*(0.007)*	*(1.144)*
Korea	79.346	0.075	0.000
	(14.584)	*(0.006)*	*(0.000)*
Malaysia	17.703	0.028	0.000
	(4.871)	*(0.002)*	*(0.000)*
Mexico	33.105	0.032	0.000
	(6.431)	*(0.003)*	*(0.000)*
Poland	68.183	0.079	0.312
	(15.238)	*(0.007)*	*(0.065)*
Russia	62.157	0.110	0.526
	(19.726)	*(0.009)*	*(0.160)*
SouthAfrica	36.599	0.055	0.254
	(8.902)	*(0.005)*	*(0.063)*

Notes: *Italic* numbers in parentheses denote the standard errors of the time-varying Linear Market Model hyperparameter estimates via KFMR.

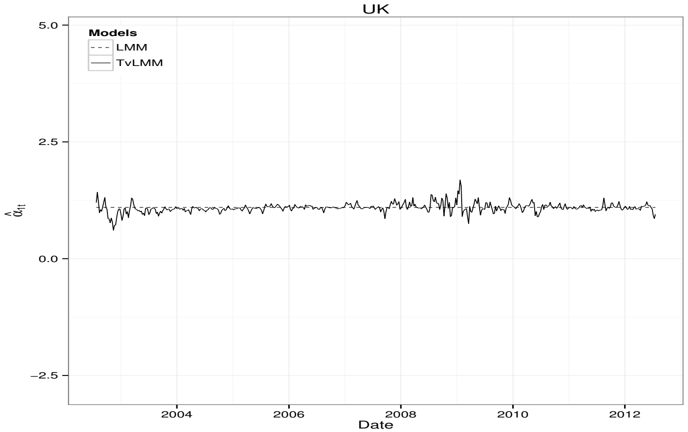
Table 5.10: Time-varying Linear Market Model state parameter estimates (standard errors) via KFMR.

Market	$\hat{\kappa}_i$	$\hat{\hat{\alpha}}_{1it}$	*Range*
Developed			
France	0.000	1.337	(0.651;1.962)
	(0.000)	*(0.043)*	
Germany	0.000	1.386	(0.302;2.561)
	(0.000)	*(0.050)*	
Italy	-0.002	1.272	(0.462;2.376)
	(0.000)	*(0.054)*	
Japan	0.000	0.741	(-0.288;1.773)
	(0.000)	*(0.036)*	
Norway	0.002	1.380	(0.281;2.772)
	(0.000)	*(0.104)*	
Sweden	0.001	1.456	(0.499;2.310)
	(0.000)	*(0.062)*	
Switzerland	0.000	0.944	(0.389;1.463)
	(0.000)	*(0.031)*	
UK	0.000	1.088	(0.608;1.686)
	(0.000)	*(0.034)*	
USA	0.000	0.910	(0.494;1.374)
	(0.000)	*(0.016)*	
Emerging			
Brazil	0.004	1.487	(-2.722;3.382)
	(0.000)	*(0.142)*	
Chile	0.003	0.862	(-0.323;1.913)
	(0.000)	*(0.046)*	
India	0.002	1.120	(0.148;2.350)
	(0.000)	*(0.177)*	
Korea	0.001	1.199	(-1.833;4.408)
	(0.000)	*(0.094)*	
Malaysia	0.002	0.582	(-0.350;1.653)
	(0.000)	*(0.025)*	
Mexico	0.003	1.167	(-0.047;2.645)
	(0.000)	*(0.060)*	
Poland	0.002	1.405	(-0.565;4.748)
	(0.000)	*(0.125)*	
Russia	0.002	1.367	(-0.333;4.230)
	(0.000)	*(0.152)*	
SouthAfrica	0.002	1.238	(-0.162;3.794)
	(0.000)	*(0.083)*	

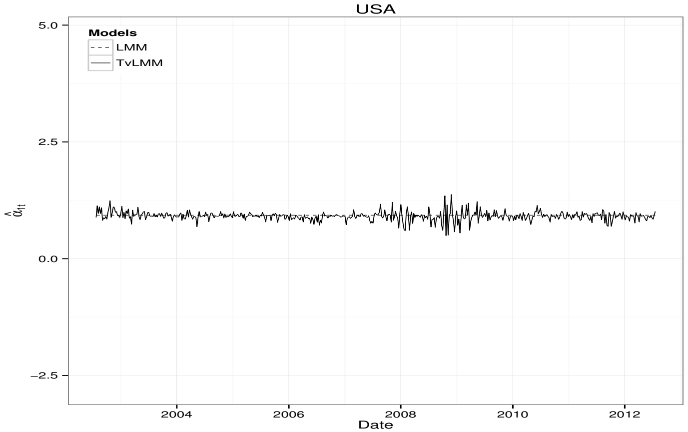
Notes: Range displays the range of $\hat{\beta}_{imt} = \hat{\alpha}_{1it}$. *Italic* numbers in parentheses denote the standard errors of the time-varying Linear Market Model state parameter estimates via KFMR.

(R_{ft}) is subtracted before estimation (see Campbell et al. (1997)). Hence, $\hat{\kappa}_i$ can be assumed to be zero in the Two-Moment CAPM in many studies (e.g. Mergner and Bulla (2008) and Choudhry and Wu (2009)) and in the previous chapter, but not here. The mean of the time-varying systematic covariance $\hat{\alpha}_{1it}$ for all of the 18 global markets is positive, and is close to 1, with a standard error close to 0.06. Note that a systematic covariance ($\hat{\alpha}_{1i}$) value of 1 means that the stock market moves in step with the MSCI World market portfolio. A value of $\hat{\alpha}_{1i}$ less than 1 means that the stock market is less volatile than the MSCI World market portfolio, whereas $\hat{\alpha}_{1i}$ greater than 1 indicates that the stock market is more volatile than the MSCI World market portfolio. The KFMR provides a wider range of time-varying systematic covariance ($\hat{\alpha}_{1it}$) risk in the emerging markets than the developed markets. This suggests that the relationship between excess returns in emerging markets and the MSCI World market portfolio as a whole is less consistent than the relationship between excess returns in developed markets and the MSCI World market portfolio as a whole.

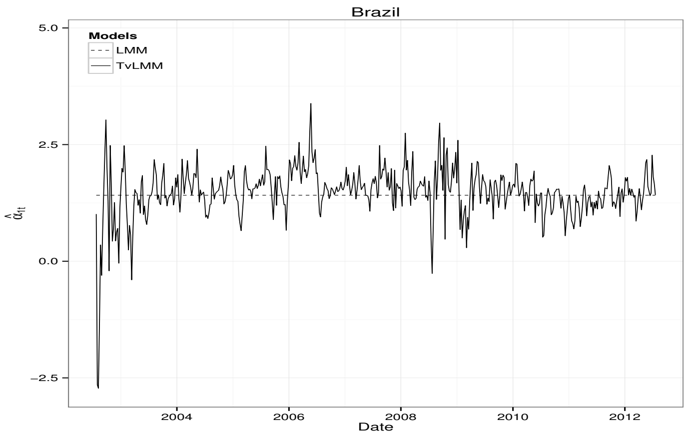
Figure 5.7 also displays the time-varying systematic covariance ($\hat{\alpha}_{1it}$) series of 2 developed markets (UK and USA) and 2 emerging markets (Brazil and Russia), respectively. To save space the remaining markets are displayed in a smaller size in Figures 5.8 and 5.9. It can be clearly seen that the time-varying systematic covariance ($\hat{\alpha}_{1it}$) series fluctuate about the equivalent Linear Market Model estimates ($\hat{\alpha}_{1i}$) in all 18 global markets. The time-varying systematic covariance ($\hat{\alpha}_{1it}$) of the emerging markets is more volatile than that of the developed markets. Also, in the presence of extreme events around October 6-10, 2008 due to the financial crisis, the time-varying systematic covariance ($\hat{\alpha}_{1it}$) is potentially more volatile in the emerging markets (Brazil and UK) than the developed markets (UK and USA). In addition, it can be clearly seen that India has qualitatively different behaviour than the rest as $\{\alpha_{1it}\}$ varies smoothly over time as $\hat{\phi}_{1i} \approx 1$. For the rest it is more random as $0 \leq \hat{\phi}_{1i} < 1$.



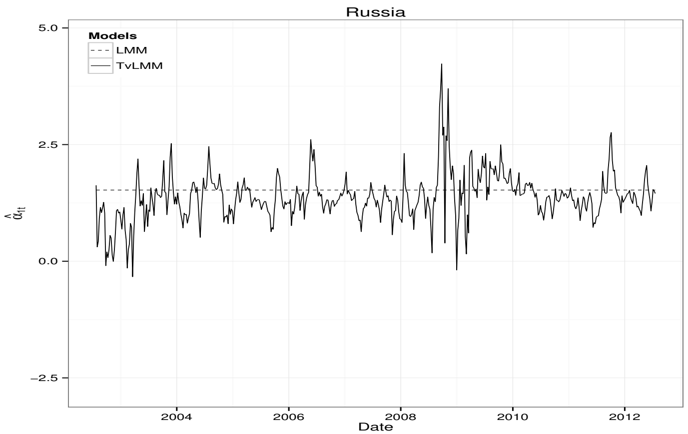
(a) UK



(b) USA

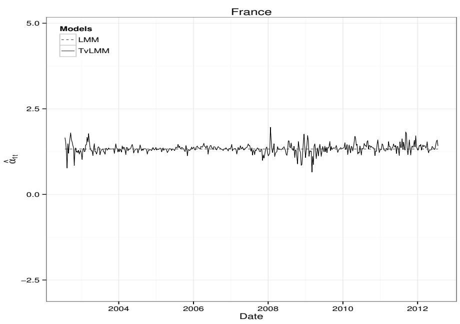


(c) Brazil

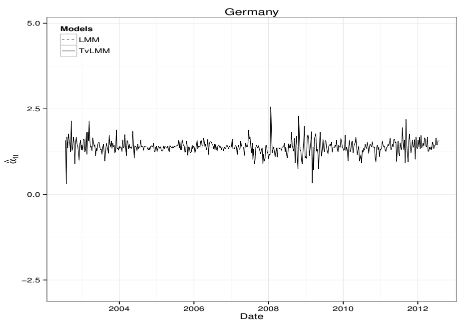


(d) Russia

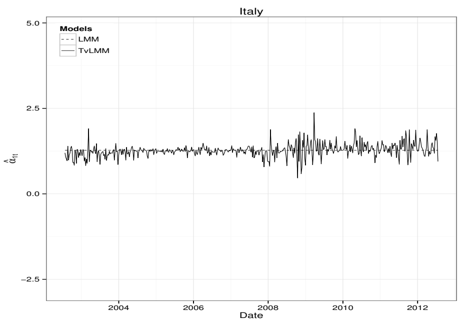
Figure 5.7: The estimated $\hat{\alpha}_{1it}$ plots of 2 developed (UK and USA) and 2 emerging (Brazil and Russia) markets.



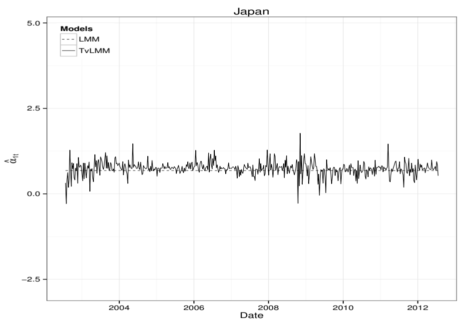
(a) France



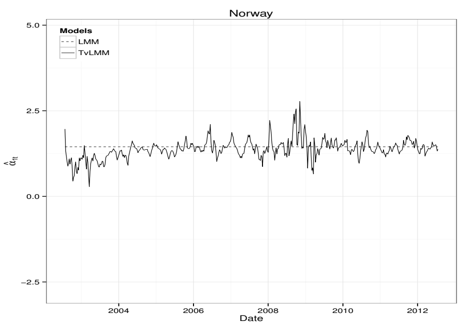
(b) Germany



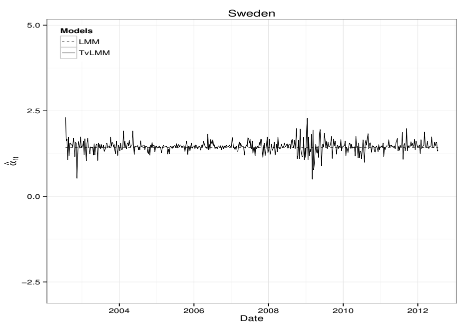
(c) Italy



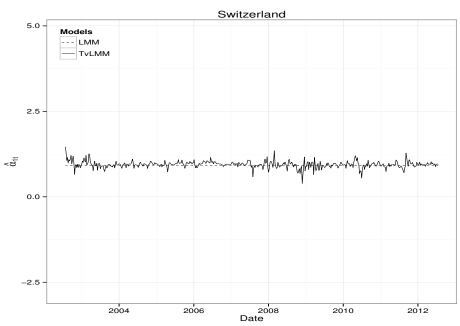
(d) Japan



(e) Norway

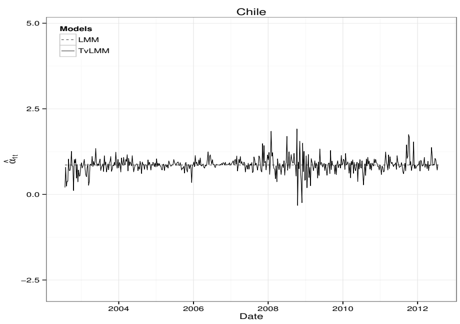


(f) Sweden

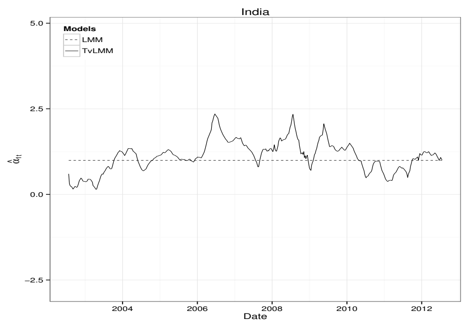


(g) Switzerland

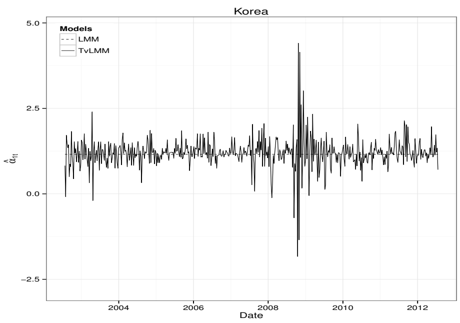
Figure 5.8: The estimated $\hat{\alpha}_{1it}$ plots of 7 developed markets.



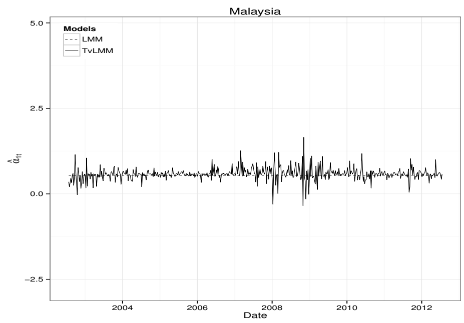
(a) Chile



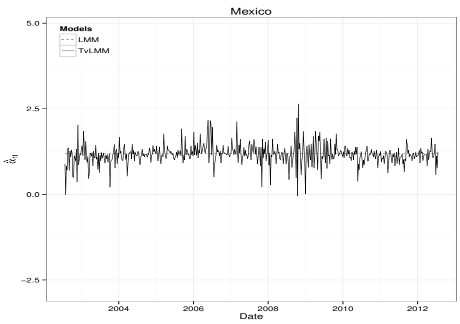
(b) India



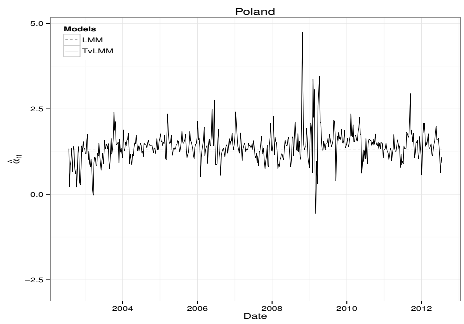
(c) Korea



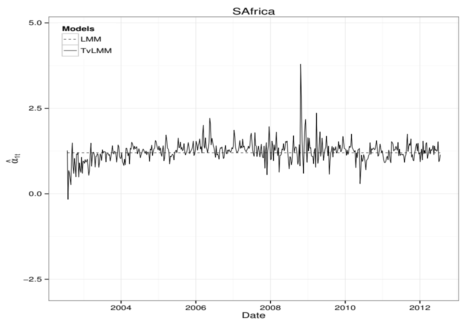
(d) Malaysia



(e) Mexico



(f) Poland



(g) South Africa

Figure 5.9: The estimated $\hat{\alpha}_{1it}$ plots of 7 emerging markets.

5.5 Comparison of Models Before and After the October 2008 Financial Crisis

A seperate comparison between the same models, for the periods before and after the October 2008 financial crisis, is made in sections 5.5.1 and 5.5.2 to investigate the effect of the crisis whilst modelling stock market returns. The comparison is made in terms of overall measures of model fit, using the Akaike Information Criterion (*AIC*), the Bayesian Information Criterion (*BIC*) and *Adjusted* R^2.

5.5.1 Model Fit Before the October 2008 Financial Crisis

The model fit results for the period from July 2002 to before the October 2008 financial crisis are given in Tables 5.11, 5.12 and 5.13, which respectively display the *AIC*, *BIC* and *Adjusted* R^2.

Table 5.11: *AIC* values for all models & markets before October 2008.

	AIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-1889.16	-1889.45	-1903.04	-1903.46	**-1907.17**	-1905.89	-1905.07
Germany	-1773.67	-1778.43	-1778.26	-1779.58	-1831.01	**-1842.46**	-1834.88
Italy	-1824.52	-1828.59	-1826.78	-1830.50	**-1836.04**	-1832.24	-1829.35
Japan	-1573.52	-1577.96	-1598.63	**-1607.15**	-1603.25	-1606.41	-1603.37
Norway	-1362.55	-1382.31	-1380.39	-1382.69	-1393.44	**-1400.08**	-1394.30
Sweden	-1674.13	-1677.62	-1685.41	**-1690.10**	-1687.17	-1683.66	-1683.10
Swit.	-1823.81	-1822.61	-1832.77	**-1835.43**	-1831.47	-1827.52	-1826.41
UK	-1905.57	-1903.93	-1903.12	-1917.45	**-1920.46**	-1914.51	-1911.68
USA	-2247.78	-2258.35	-2259.06	-2258.72	**-2267.24**	-2266.58	-2264.01
Emerging							
Brazil	-1149.20	-1189.76	-1202.31	-1212.53	-1281.11	**-1283.82**	-1281.81
Chile	-1499.76	-1502.93	-1518.17	-1515.42	**-1518.82**	-1516.38	-1513.46
India	-1251.66	-1256.96	-1267.28	-1267.40	**-1297.10**	-1293.91	-1292.98
Korea	-1350.46	-1355.16	-1366.39	**-1368.13**	-1357.08	-1353.04	-1355.65
Malaysia	-1623.05	-1633.95	**-1641.67**	-1639.86	-1641.28	-1639.42	-1641.51
Mexico	-1508.42	-1520.91	-1533.71	-1535.39	-1555.29	**-1561.70**	-1556.77
Poland	-1318.39	-1327.18	-1329.08	-1331.56	-1331.72	**-1332.88**	-1329.45
Russia	-1137.42	-1151.23	-1150.44	-1153.17	**-1186.18**	-1184.67	-1178.54
S.Africa	-1414.34	-1430.55	-1431.98	-1434.06	-1447.90	**-1454.81**	-1446.74

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest *AIC*.

Table 5.11 shows that in the period before the October 2008 financial crisis, the lowest AIC value comes from the time-varying Linear Market (TvLMM) and the time-varying Quadratic Market Models (TvQMM) via KFMR for all developed markets and emerging markets, with the exceptions being Japan, Sweden, Switzerland and Korea, where the generalized additive model (GAM) has the lowest AIC value, and Malaysia, where the Cubic Market Model (CMM) has the lowest AIC value. In addition, the GAM outperforms the Higher order DGPs in general across most of the 18 global markets, with the exception of the USA, Chile and Malaysia. The Linear Market Model (LMM) also outperforms the Higher order DGPs in the UK.

Table 5.12: BIC values for all models & markets before October 2008.

	BIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-1877.83	-1874.34	-1884.15	-1881.57	**-1888.29**	-1875.66	-1863.51
Germany	-1762.34	-1763.32	-1759.37	-1762.32	**-1812.12**	-1811.84	-1793.32
Italy	-1813.19	-1813.48	-1807.89	-1796.20	**-1817.15**	-1802.03	-1787.80
Japan	-1562.18	-1562.85	-1579.74	-1571.74	**-1584.36**	-1576.19	-1561.82
Norway	-1351.22	-1367.20	-1361.50	-1350.37	**-1374.55**	-1369.86	-1352.74
Sweden	-1662.80	-1662.51	-1666.53	-1658.57	**-1668.28**	-1653.43	-1641.55
Swit.	-1812.48	-1807.49	**-1813.88**	-1809.91	-1812.58	-1797.30	-1784.86
UK	-1894.24	-1888.82	-1884.23	-1879.54	**-1901.57**	-1884.29	-1870.13
USA	-2236.45	-2243.24	-2240.17	-2240.10	**-2248.35**	-2236.36	-2222.45
Emerging							
Brazil	-1137.87	-1174.65	-1183.42	-1174.32	**-1262.22**	-1253.59	-1240.26
Chile	-1488.43	-1487.82	-1499.28	-1489.54	**-1499.94**	-1486.16	-1471.90
India	-1240.33	-1241.85	-1248.39	-1239.24	**-1278.21**	-1263.69	-1251.42
Korea	-1339.13	-1340.05	**-1347.50**	-1337.09	-1338.19	-1322.82	-1314.10
Malaysia	-1611.72	-1618.83	**-1622.78**	-1618.64	-1622.39	-1609.20	-1599.96
Mexico	-1497.09	-1505.80	-1514.82	-1512.41	**-1536.40**	-1531.48	-1515.22
Poland	-1307.06	-1312.07	-1310.19	-1311.04	**-1312.84**	-1302.66	-1287.90
Russia	-1126.09	-1136.12	-1131.56	-1119.65	**-1167.29**	-1154.45	-1136.98
S.Africa	-1403.00	-1415.44	-1413.09	-1412.66	**-1429.01**	-1424.59	-1405.18

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest *BIC*.

Table 5.12, on the other hand, illustrates that, based on the lowest BIC, the time-varying Linear Market Model (TvLMM) via KFMR outperforms all of the other models during the period before the October 2008 financial crisis for all developed and emerging markets, with the exception of Switzerland, Korea and

Malaysia, where the Cubic Market Model (CMM) is more appropriate. Within the time-varying Higher order DGPs, the time-varying Quadratic Market Model (TvQMM) outperforms the time-varying Cubic Market Model (TvCMM) for all 18 global markets; suggesting that in this case the time-varying Quadratic Market Model is preferable to the time-varying Cubic Market Model for all 18 global markets. Moreover, it has been shown that the Higher order DGPs outperform the GAM for all developed and emerging markets. This is because the *BIC* has a larger penalty than the *AIC*, and thus chooses simpler models. These results show that the time-varying Linear Market Model (TvLMM) is preferable to other models, in that the additional complexity of other models does not greatly improve their model fit performance in relation to the time-varying Linear Market Model.

Table 5.13: *Adjusted* R^2 values for all models & markets before October 2008.

	Adjusted R^2						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	0.829	0.829	0.837	0.838	**0.889**	0.876	0.886
Germany	0.785	0.789	0.789	0.790	**0.899**	0.874	0.879
Italy	0.709	0.713	0.712	0.719	**0.808**	0.804	0.802
Japan	0.344	0.355	0.397	0.420	**0.575**	0.545	0.527
Norway	0.431	0.466	0.464	0.474	0.546	**0.567**	0.563
Sweden	0.754	0.758	0.764	0.770	**0.828**	0.824	0.806
Swit.	0.716	0.716	0.726	0.730	**0.791**	0.787	0.776
UK	0.764	0.764	0.764	0.777	0.814	**0.815**	0.814
USA	0.874	0.878	0.879	0.879	**0.923**	0.901	0.898
Emerging							
Brazil	0.294	0.379	0.405	0.432	**0.790**	0.772	0.775
Chile	0.334	0.343	0.375	0.373	**0.512**	0.502	0.486
India	0.235	0.250	0.276	0.281	0.559	0.547	**0.567**
Korea	0.332	0.343	0.368	0.377	**0.506**	0.488	0.433
Malaysia	0.245	0.272	0.292	0.289	**0.560**	0.534	0.467
Mexico	0.496	0.517	0.537	0.541	**0.678**	0.669	0.663
Poland	0.325	0.345	0.351	0.357	0.414	**0.424**	0.412
Russia	0.227	0.262	0.262	0.277	**0.532**	0.513	0.514
S.Africa	0.407	0.438	0.442	0.447	0.519	0.536	**0.548**

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the highest *Adjusted* R^2.

The results for *Adjusted* R^2 are given in Table 5.13. Again these show that the time-varying Linear Market Model (TvLMM) via KFMR performs better than the time-varying Higher order DGPs for all 18 global markets, with the exceptions being that of Norway, the UK and Poland, where the time-varying Quadratic Market Model (TvQMM) is better, as well as India and South Africa, where the time-varying Cubic Market Model (TvCMM) is better. Nevertheless, use of these models does not result in improvements in model fit over the time-varying Linear Market Model. For example, the time-varying Linear Market Model improves on the time-varying Quadratic Market Model in terms of *Adjusted* R^2 by on average 0.9% for both the developed and emerging markets. The time-varying Linear Market Model also improves on the time-varying Cubic Market Model in terms of *Adjusted* R^2 by on average 1.4% for the developed markets and 2.3% for the emerging markets. These results confirm what is reported in Tables 5.11 and 5.12, namely, that the time-varying Linear Market Model (TvLMM) outperforms all other time-varying DGPs. This confirms that the additional complexity of the time-varying Higher order DGPs does not improve the model fit performance of the time-varying Linear Market Model during the period before the October 2008 financial crisis.

The time-varying Linear Market Model (TvLMM) provides a much better performance during this period than the Higher order DGPs and the generalized additive model (GAM). The time-varying Linear Market Model improves on the Linear Market Model (LMM) in terms of *Adjusted* R^2 by on average 9.6% for the developed markets and 24.2% for the emerging markets, suggesting that, during the period before the October 2008 financial crisis, the emerging markets were more unstable than the developed markets. In addition, the GAM outperforms all DGPs for all 18 global markets, although the improvements in model fit are not substantial. For example, the GAM improves on the Linear Market Model in terms of *Adjusted* R^2 by on average only 2.1% for the developed markets and 5.3% for the emerging markets. The average increase in *Adjusted* R^2 from the Linear Market Model to the Cubic Market Model (CMM) is 1.4% for the developed markets and 4.6% for the emerging markets. The average increase in *Adjusted* R^2 from the Linear Market Model to the Quadratic Market Model (QMM) is 0.7% for the developed markets and 2.8% for the emerging markets. As seen in Tables 5.11 and 5.12, these results confirm that, in the period before the October

2008 financial crisis, the time-varying Linear Market Model offers a substantial improvement in model fit to the Linear Market Model, but the other non-linear DGPs do not.

5.5.2 Model Fit After the October 2008 Financial Crisis

The model fit for the period from after the October 2008 financial crisis to July 2012 results are given in Tables 5.14, 5.15 and 5.16, which respectively display the *AIC*, *BIC* and *Adjusted* R^2.

Table 5.14: *AIC* values for all models & markets after October 2008.

	AIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-1068.73	-1068.98	-1068.81	-1071.16	**-1085.15**	-1080.62	-1075.70
Germany	-1024.33	-1022.78	-1021.60	-1024.33	**-1048.74**	-1042.78	-1037.39
Italy	-928.17	-929.79	-928.30	**-929.95**	-928.33	-924.35	-918.69
Japan	-955.82	-954.68	-955.97	-956.01	**-957.48**	-953.08	-949.25
Norway	-878.52	-885.72	-884.10	-891.58	**-896.71**	-893.37	-887.93
Sweden	-908.17	-906.31	-907.21	-908.17	**-920.59**	-915.00	-909.98
Swit.	-1076.68	-1075.17	**-1078.55**	-1077.87	-1074.94	-1069.57	-1068.38
UK	-1114.71	-1113.97	-1111.98	-1114.71	**-1118.06**	-1112.42	-1106.35
USA	-1316.31	-1314.32	-1318.88	-1320.68	**-1330.70**	-1324.73	-1322.32
Emerging							
Brazil	-872.91	-870.92	-870.18	-872.91	**-876.30**	-870.30	-865.62
Chile	-877.85	-879.19	-877.76	-880.39	**-889.21**	-883.91	-879.45
India	-780.97	-780.45	-781.28	-781.58	**-788.75**	-786.02	-780.22
Korea	-785.30	-783.67	-783.21	-785.30	**-789.66**	-783.71	-779.02
Malaysia	-1015.19	-1013.84	-1015.49	-1015.60	**-1019.80**	-1014.10	-1010.21
Mexico	-972.24	-971.82	-970.13	-972.25	**-977.82**	-972.23	-966.44
Poland	-721.17	-720.08	-720.76	-726.05	**-740.00**	-734.32	-729.92
Russia	-736.71	-739.21	-738.44	-739.82	**-746.23**	-744.30	-738.94
S.Africa	-877.69	-876.85	-874.99	**-877.82**	-876.40	-871.76	-865.76

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest *AIC*.

Table 5.14 shows that, based on the lowest *AIC*, the time-varying Linear Market Model (TvLMM) via KFMR is the most appropriate model for all 18 global markets, with the exception of Italy and South Africa, where the generalized additive model (GAM) has the lowest *AIC* value, and of Switzerland, where the Cubic Market Model (CMM) has the lowest *AIC* value. It can be clearly seen that the time-varying Quadratic Market Model (TvQMM) outperforms the time-varying Cubic Market Model (TvCMM) for all 18 global markets, and is thus preferable to the latter. In addition, the GAM outperforms the time-varying Cubic Market Model for 5 of the developed markets, including Italy, Japan, Norway, Switzerland and the UK, as well as all of the emerging markets except for Poland. This suggests that the additional complexity of the time-varying Cubic Market Model via KFMR does not improve the model fit in terms of *AIC*. The GAM also outperforms the Higher order DGPs for all 18 global markets apart from Switzerland, where the Cubic Market Model is preferred. It can be clearly seen that the Linear Market Model (LMM) outperforms the Higher order DGPs in Germany, Sweden, the UK, Brazil, Korea, Mexico, Poland and South Africa, suggesting that the additional complexity of the Higher order DGPs does not greatly improve model fit in terms of *AIC* during the period after the October 2008 financial crisis.

Table 5.15: *BIC* values for all models & markets after October 2008.

	BIC						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	-1058.94	-1055.93	-1052.50	-1053.61	**-1068.84**	-1054.52	-1039.79
Germany	-1014.54	-1009.73	-1005.29	-1014.54	**-1032.43**	-1016.67	-1001.50
Italy	**-918.38**	-916.74	-911.99	-917.08	-912.02	-898.25	-882.80
Japan	**-946.03**	-941.63	-939.66	-943.75	-941.17	-926.98	-913.36
Norway	-868.73	-872.67	-867.79	-859.57	**-880.40**	-867.27	-852.04
Sweden	-898.38	-893.26	-890.90	-898.38	**-904.28**	-888.90	-874.09
Swit.	**-1066.89**	-1062.12	-1062.24	-1061.85	-1058.63	-1043.47	-1032.46
UK	**-1104.92**	-1100.92	-1095.67	**-1104.92**	-1101.75	-1086.31	-1070.46
USA	-1306.53	-1301.27	-1302.57	-1299.09	**-1314.39**	-1298.62	-1286.43
Emerging							
Brazil	**-863.12**	-857.87	-853.87	**-863.12**	-859.99	-844.20	-829.73
Chile	-868.07	-866.14	-861.45	-865.38	**-872.89**	-857.81	-843.56
India	-771.18	-767.40	-764.96	-767.33	**-772.43**	-759.92	-744.33
Korea	**-775.52**	-770.62	-766.90	**-775.52**	-773.35	-757.61	-743.13
Malaysia	**-1005.40**	-1000.79	-999.18	-999.95	-1003.48	-988.00	-974.32
Mexico	**-962.46**	-958.77	-953.82	-962.21	-961.51	-946.14	-930.55
Poland	-711.38	-707.03	-704.45	-693.42	**-723.68**	-708.21	-694.03
Russia	-726.92	-726.16	-722.12	-718.98	**-729.92**	-718.20	-703.05
S.Africa	**-867.91**	-863.80	-858.68	-866.87	-860.09	-845.65	-829.87

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the lowest *BIC*.

Table 5.15 shows that the lowest *BIC* value comes from the time-varying Linear Market Model (TvLMM) via KFMR during the period after the October 2008 financial crisis for 5 of the developed markets, including France, Germany, Norway, Sweden and the USA, as well as 4 of the emerging markets, including Chile, India, Poland and Russia. It can also be seen that the Linear Market Model (LMM) outperforms the time-varying Higher order DGPs and the Higher order DGPs for 4 of the developed markets including Italy, Japan, Switzerland and the UK, as well as 5 of the emerging markets, including Brazil, Korea, Malaysia, Mexico and South Africa. This is because the *BIC* has a larger penalty than the *AIC*, and thus chooses simpler models. These results suggest that the additional complexity of these models does not greatly improve model fit compared to the Linear Market Model in terms of *BIC* during the period after the October 2008 financial crisis.

Table 5.16: *Adjusted* R^2 values for all models & markets after October 2008.

	Adjusted R^2						
Model	LMM	QMM	CMM	GAM	TvLMM	TvQMM	TvCMM
Developed							
France	0.884	0.885	0.885	0.887	**0.944**	0.942	0.942
Germany	0.860	0.859	0.859	0.860	**0.944**	0.943	0.943
Italy	0.814	0.816	0.816	0.817	**0.863**	0.859	0.858
Japan	0.453	0.452	0.459	0.455	**0.624**	0.621	0.597
Norway	0.792	0.800	0.800	0.812	**0.891**	0.883	0.878
Sweden	0.782	0.781	0.783	0.782	**0.877**	0.866	0.875
Swit.	0.760	0.760	0.765	0.764	**0.809**	0.808	0.770
UK	0.872	0.872	0.871	0.872	**0.922**	0.921	0.920
USA	0.929	0.929	0.931	0.932	**0.966**	0.965	0.963
Emerging							
Brazil	0.734	0.733	0.733	0.734	**0.792**	0.790	0.790
Chile	0.468	0.475	0.474	0.480	0.650	0.695	**0.698**
India	0.484	0.486	0.490	0.490	0.588	**0.600**	0.593
Korea	0.564	0.562	0.564	0.564	**0.739**	0.737	0.733
Malaysia	0.463	0.462	0.469	0.469	**0.690**	0.686	0.677
Mexico	0.752	0.752	0.752	0.752	**0.863**	0.861	0.861
Poland	0.618	0.617	0.621	0.640	0.841	0.841	**0.842**
Russia	0.658	0.664	0.665	0.669	**0.785**	0.761	0.758
S.Africa	0.697	0.697	0.696	0.698	0.769	**0.774**	0.772

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa. **Bold** displays the best market pricing model for each market in terms of the highest *Adjusted* R^2.

The results for *Adjusted* R^2 are given in Table 5.16. This again shows that the time-varying Linear Market Model (TvLMM) via KFMR provides a better performance than the time-varying Higher order DGPs during the period after the October 2008 financial crisis for all developed and emerging markets, with the exceptions being Chile and Poland, where the time-varying Cubic Market Model (TvCMM) performs better, as well as India and South Africa, where the time-varying Quadratic Market Model (TvQMM) achieves better results. The improvements in model fit to the time-varying Linear Market Model, however, are not substantial. For example, the time-varying Linear Market Model improves on the time-varying Quadratic Market Model in terms of *Adjusted* R^2 by on average 0.4% for the developed markets; however, the time-varying Linear Market Model reduces on the time-varying Quadratic Market Model in terms of *Adjusted* R^2 by on average 0.3% for the emerging markets. In addition, whilst the time-varying Linear Market Model improves on the time-varying Cubic Market Model in terms of *Adjusted* R^2 by on average 1% for the developed markets, the time-varying Linear Market Model worsens on the time-varying Cubic Market Model in terms of *Adjusted* R^2 by on average 0.1% for the emerging markets. These results may be due to the fact that the emerging markets were more unstable than the developed markets during this period, and might have become more integrated with the October 2008 financial crisis than the developed markets. These results confirm what is seen in Tables 5.14 and 5.15, namely that adding the additional complexity of the time-varying Higher order DGPs provides a slight improvement in model fit in comparison to the time-varying Linear Market Model during the period after the October 2008 financial crisis in the emerging markets, but not in the developed markets.

The time-varying Linear Market Model (TvLMM) via KFMR again provides a much better performance than the Higher order DGPs and the generalized additive model (GAM) for this period. The time-varying Linear Market Model improves on the Linear Market Model (LMM) in terms of *Adjusted* R^2 by on average 7.7% for the developed markets and 14.2% for the emerging markets. It can be seen that the GAM outperforms the Higher order DGPs for all 18 global markets, but that the improvements in model fit are not substantial. For example, the average increase in *Adjusted* R^2 from the Cubic Market Model

(CMM) to the GAM is 0.1% for the developed markets and 0.4% for the emerging markets. Furthermore, the GAM improves on the Linear Market Model in terms of *Adjusted* R^2 by on average only 0.4% for the developed markets and 0.6% for the emerging markets. The average increase in *Adjusted* R^2 from the Linear Market Model to the Cubic Market Model is 0.3% for both the developed and emerging markets. These results confirm what is seen in Tables 5.14 and 5.15, namely, that in the period under discussion, the time-varying Linear Market Model offers a substantial improvement in model fit to the Linear Market Model, but that the other non-linear DGPs do not.

5.6 Conclusion

The focus of this chapter has been on assessing the appropriateness of the Linear Market Model (consistent with the Two-Moment CAPM) by comparing it to Higher order DGPs, consistent with their equivalent Higher-Moment CAPMs, a generalized additive model (GAM), the time-varying Linear Market Model and the time-varying Higher order DGPs via KFMR when applied to the 18 global markets during the three different time periods (July 2002-July 2012; July 2002-before October 2008; and after October 2008-July 2012) for the purpose of investigating the effect of the October 2008 financial crisis when modelling stock market returns.

The appropriateness of the models was assessed by overall measures of model fit using *AIC*, *BIC* and *Adjusted* R^2, residual diagnostics, and by a graphical summary of the fitted models to the data over the entire period of July 2002-2012. The Linear Market Model (LMM) performs worse under both metrics. The Cubic Market Model (CMM) provides a slight improvement on the Linear Market Model in terms of *Adjusted* R^2, by on average 0.4% for the developed markets and 2% for the emerging markets. Using a Quadratic Market Model (QMM), the *Adjusted* R^2 decreases compared to the Linear Market Model for Germany, India and Malaysia. In addition, the GAM outperforms the Higher order DGPs in all 18 global markets, but the GAM improves on the Linear Market Model in terms of *Adjusted* R^2 by on average only 1% for the developed markets and 3.9% for the emerging markets. The time-varying Linear Market Model (TvLMM) via KFMR provides a much better performance than any of the non-linear DGPs. Furthermore, it improves on the Linear Market Model in terms of *Adjusted* R^2 by on average 10.4% for the developed markets and 25.3% for the emerging markets. The time-varying Linear Market Model also provides a better performance than the time-varying Higher order DGPs, but the improvements in model fit are not substantial. The time-varying Linear Market Model improves on the time-varying Quadratic Market Model (TvQMM) in terms of *Adjusted* R^2 by on average 0.3% for the developed markets and 0.6% for the emerging markets. The time-varying Linear Market Model also improves on the time-varying Cubic Market Model (TvCMM) in terms of *Adjusted* R^2 by on average 0.5% for the developed markets and 1% for the emerging markets. These results confirm that the additional complexity of the time-varying Higher order DGPs does not provide an improvement in model fit compared to the time-varying Linear Market Model for both the developed and emerging markets over the entire period from July 2002 to July 2012.

The model fit results for the periods before and after the October 2008 financial crisis also confirm that the Linear Market Model performs worse in both cases. The GAM outperforms the Higher order DGPs in terms of average *Adjusted* R^2 for both the developed and emerging markets, but it provides only a slight improvement on the Linear Market Model. For example, the GAM improves on the Linear Market Model in terms of *Adjusted* R^2, by on average 2.1% for the developed markets and 5.3% for the emerging markets during the period before the October 2008 financial crisis, and an average 0.4% for the developed markets and 0.6% for the emerging markets in the post-crisis period. The time-varying Linear Market Model provides overall a much better performance than any of the other models in the periods both before and after the crisis for both the developed and emerging markets; the exception being the emerging markets in the period after the crisis, where the time-varying Quadratic Market Model performs better in terms of average *Adjusted* R^2. However, in this case the improvements in model fit to the time-varying Linear Market Model are not substantial. For example, the time-varying Quadratic Market Model improves on the time-varying Linear

Market Model in terms of *Adjusted* R^2 by on average 0.3% for the emerging markets. It can be seen that the improvements in model fit in terms of average *Adjusted* R^2 from the Linear Market Model to the time-varying Linear Market Model for the developed markets (9.6%; and 7.7%, respectively), are less than the corresponding improvements for the emerging markets (24.2%; and 14.2%, respectively) during these periods. These results may be due to the fact that the developed markets were more stable than the emerging markets during the periods both before and after the October 2008 financial crisis.

These results support the pricing of the time-varying systematic covariance risk rather than the other time-varying systematic risk measures, time-varying systematic skewness and time-varying systematic kurtosis as well as the time-invariant systematic risk measures, systematic covariance, systematic skewness and systematic kurtosis in the CAPMs. They also confirm the instability of the systematic covariance risk mentioned in the existing literature and the previous chapter with regard to the Two-Moment CAPM.

Note that residual diagnostics are discussed in section 5.4.2. These assumptions are that the residuals are normally distributed, independent (no autocorrelation) and have constant variance (homoskedasticity). When these assumptions are violated, the state space models can be affected but appropriate extensions of the state space model were discussed in section 4.6.

Chapter 6

Multivariate State Space Modelling

6.1 Introduction

The Two-Moment Capital Asset Pricing Model (CAPM), devised by Sharpe (1964), Lintner (1965) and Mossin (1966), remains a well known and still widely used pricing model. The model implies a linear relationship between the expected return on a countries stock market and its systematic covariance (*beta*) risk, which is constant over time. However, as discussed in the previous chapters, there is substantial empirical evidence to suggest that systematic covariance risk varies over time, possibly in response to economic factors, such as unemployment rate, credit score, etc.. In Chapter 5, the appropriateness of the unconditional and conditional Two-Moment CAPMs for developed and emerging stock markets was assessed, and the results suggested that a time-varying Linear Market Model (consistent with the conditional Two-Moment CAPM) provides a much better performance when explaining stock market excess returns than a simpler Linear Market Model whose systematic covariance risk is constant over time. However, the performance of these models was evaluated in a univariate context separately for each stock market, which does not utilise the correlation structure among different countries' stock markets in the estimation process.

The correlation in countries' stock markets is understood to be a consequence of economic and financial integration. Research by Yavas (2007) reveals that increasing levels of international trade and international financial transactions have influenced the correlation between national economies worldwide in recent decades. This economic integration has contributed to an increased level of correlation between stock markets and heightened their integration.

According to the CAPM and portfolio theory (Markowitz (1952)), the likely presence of correlations between various stock markets is important for generating a well diversified portfolio to reduce overall risk. This encouraged financial researchers to explore dependencies between stock markets in a multivariate Two-Moment CAPM. For example, prominent papers such as Gibbons et al. (1989), Mackinlay and Richardson (1991) and Hansen and Jagannathan (1997) have considered analysing the assets pricing data model in a multivariate context with time invariant systematic covariance risk. This chapter considers this multivariate context across several countries' stock markets simultaneously, and extends the above literature to consider time-varying systematic covariance risks. Our overall approach is to extend the time-varying systematic covariance risk models used in Chapter 5 to the multivariate domain, which allows the between stock market correlation to be utilised in the estimation process. A multivariate state space approach, which is the first such approach in this literature, is taken, and considering the superiority of the performance of the mean reverting specification in Chapter 4 and Chapter 5, this is the only model considered here. (The random walk and random coefficients models are special cases and are not considered explicitly.). Note that the direction of the stock market co-movement can be measured using various methodologies such as multivariate GARCH (e.g. Chiang et al. (2007), Wang and Moore (2008) and Saleem (2009)) and vector auto-regression (VAR) (e.g. Liu et al. (1998), Chang and Nieh (2001) and Jayasuriya (2011)) in the financial literature.

The main purpose of this chapter is to model and forecast time-varying systematic covariance (*beta*) risks based on a multivariate state space form of the time-varying Linear Market Model (consistent with conditional Two-Moment CAPM) using a Kalman Filter Mean Reverting model, and to see if this outperforms a univariate approach. Both in-sample modelling and out-of-sample forecasting procedures are considered, and are used to quantify model performance. Four possible multivariate state space model formulations are considered, a comparison that is yet to be undertaken in the literature. The aim of this chapter is to compare the performance of these models to evaluate the extent to which accounting for the between country stock market correlation structure aids in modelling and forecasting country's stock market returns. The comparison is made using weekly data (as also used in Chapter 5), generated by 9 developed (France, Germany, Italy, Japan, Norway, Sweden, Switzerland, UK and USA) and 9 emerging markets (Brazil, Chile, India, Korea, Malaysia, Mexico, Poland, Russia and South Africa) over the period from July 2002 to July 2012 together with the three-month US dollar London Interbank Offered Rate (LIBOR) interest as a proxy for the risk-free rate. In all cases the Morgan Stanley Capital International (MSCI) World Index is used as a proxy for the market portfolio. The modelling and forecasting abilities of models are evaluated using two different measures of error, the Mean Square Error (MSE) and the Mean Absolute Error (MAE).

The distributional characteristics of this data set were presented in section 5.3. However, to motivate the multivariate modelling undertaken here, Table 6.1 displays the correlation coefficients between stock markets. The table shows correlations within developed (top panel), and emerging (middle panel) markets, as well as between the two market groups (bottom panel). All correlation coefficients are positive between stock markets, which means that all of the stock markets move in the same direction, up or down, instead of moving in the opposite direction. The average correlation among the developed markets (0.76) is higher than that of the emerging markets (0.60). This may be explained by volatility, which is lower in developed markets hence resulting in higher between stock market correlations. For example, researches on the Japanese and US stock market integration by Takatoshi and Wen-Ling (1993) and on the German, US and Japanese stock market integration by Yavas (2007) have indicated that volatility affects cross stock market correlation. In addition, the average correlation between the emerging markets and the developed markets is 0.61 which is about the same as the average correlation within emerging markets. These findings match those identified by previous researchers (e.g. Harvey (1995), Bekaert and Harvey (1997)) and suggest a priori that a multivariate approach will be of most

Table 6.1: Correlation matrix between countries' stock markets.

	Developed Markets								
	France	Germany	Italy	Japan	Norway	Sweden	Swit.	UK	USA
France	1.00	0.95	0.93	0.57	0.79	0.89	0.90	0.91	0.81
Germany		1.00	0.89	0.57	0.76	0.88	0.86	0.87	0.80
Italy			1.00	0.56	0.78	0.82	0.83	0.86	0.75
Japan				1.00	0.56	0.54	0.54	0.53	0.49
Norway					1.00	0.78	0.72	0.80	0.67
Sweden						1.00	0.83	0.85	0.78
Swit.							1.00	0.85	0.73
UK								1.00	0.80
USA									1.00
	Emerging Markets								
	Brazil	Chile	India	Korea	Malaysia	Mexico	Poland	Russia	S.Africa
Brazil	1.00	0.68	0.55	0.55	0.50	0.76	0.63	0.68	0.72
Chile		1.00	0.51	0.53	0.51	0.68	0.50	0.59	0.60
India			1.00	0.61	0.61	0.57	0.54	0.57	0.58
Korea				1.00	0.62	0.62	0.54	0.58	0.57
Malaysia					1.00	0.52	0.52	0.53	0.51
Mexico						1.00	0.65	0.67	0.71
Poland							1.00	0.68	0.72
Russia								1.00	0.67
S.Africa									1.00
	Brazil	Chile	India	Korea	Malaysia	Mexico	Poland	Russia	S.Africa
France	0.66	0.62	0.57	0.60	0.54	0.70	0.70	0.65	0.75
Germany	0.67	0.60	0.57	0.62	0.53	0.70	0.71	0.65	0.74
Italy	0.64	0.60	0.58	0.59	0.56	0.69	0.70	0.67	0.70
Japan	0.49	0.47	0.52	0.66	0.54	0.52	0.48	0.54	0.52
Norway	0.68	0.59	0.57	0.60	0.58	0.66	0.69	0.74	0.71
Sweden	0.64	0.61	0.56	0.62	0.52	0.70	0.69	0.64	0.70
Swit.	0.56	0.56	0.52	0.54	0.48	0.65	0.63	0.56	0.68
UK	0.65	0.63	0.58	0.58	0.57	0.70	0.69	0.65	0.75
USA	0.64	0.59	0.50	0.53	0.46	0.75	0.59	0.58	0.63

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa.

benefit in modelling developed markets.

It is worth noting that Japan has the lowest correlations within the developed markets. This implies that Japan is the least integrated with the other developed markets due perhaps to having fewer mutual trade and financial transactions. Conversely, Japan has a higher correlation with Korea (an emerging market) than any of the developed markets, which is possibly a consequence of its geographical proximity and the fact that both are affected by regional factors such as political decisions or environmental influences (e.g. earthquake, flood). These are called common border effects by Flavin et al. (2002).

The rest of this chapter is outlined as follows. Section 6.2 introduces the multivariate state space model and presents the alternative correlation structures to be compared. Section 6.3 presents the empirical results obtained from the model comparison, and section 6.4 presents further results from the best forecasting model. Section 6.5 presents the extension of the best forecasting model. Finally, section 6.6 presents our conclusions.

6.2 Methodology

The model proposed here is a multivariate state space model, which generalizes the univariate Kalman Filter Mean Reverting (KFMR) model presented in section 3.3 and applied in Chapter 5 to data on 9 developed and 9 emerging markets. The model is applied to data on the 9 developed and 9 emerging markets separately, as the two are thought to exhibit different behaviours. Thus, the models considered will be applied to a vector of 9 stock market excess returns in this chapter, but can obviously be extended to other dimensions. The response vector here is $\mathbf{R}_t - \mathbf{R}_{ft} = (R_{1t} - R_{ft}, \ldots, R_{9t} - R_{ft})^{'}$, the 9×1 vector of excess returns for developed or emerging markets in week t.

The general form of the observation and state equations for the model proposed for these data are given by

$$\mathbf{R}_t - \mathbf{R}_{ft} = \boldsymbol{\kappa} + ((R_{mt} - R_{ft}) \otimes I_9)\, \boldsymbol{\alpha}_{1t} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \backsim N(\mathbf{0}, H), \tag{6.1}$$

$$\boldsymbol{\alpha}_{1t} = \bar{\boldsymbol{\alpha}}_1 + \Phi\,(\boldsymbol{\alpha}_{1\,t-1} - \bar{\boldsymbol{\alpha}}_1) + \boldsymbol{w}_t, \quad \boldsymbol{w}_t \backsim N(\mathbf{0}, Q). \tag{6.2}$$

The 9×1 state vector at time t for the set of developed or emerging markets is denoted by $\boldsymbol{\alpha}_{1t} = (\alpha_{11t}, \ldots, \alpha_{19t})^{'}$, and quantifies the time-varying relationship between each country's stock market $(\boldsymbol{R}_t - \boldsymbol{R}_{ft})$ and the MSCI market portfolio for week t. The vector is initialised at time zero by

$$\boldsymbol{\alpha}_{10} \sim N(\boldsymbol{\mu}_{\alpha_1}, \Sigma_{\alpha_1}), \tag{6.3}$$

where the parameters of this distribution estimated from the data as part of the estimation algorithm.

The observation and state errors ($\boldsymbol{\varepsilon}_t$, $\boldsymbol{w}_t$) are assumed to be mutually independent of each other and independent in time, and are assigned multivariate Gaussian distributions with zero means and variance matrices (H, Q) respectively. Finally, Φ is a 9×9 matrix quantifying the temporal autocorrelations in the state vector within each stock market, and is typically assumed to be diagonal so that the between stock market correlations in the state vector are modelled through Q. Thus, in this formulation H captures the correlations between stock markets in the data, while Q captures these correlations in the state vector.

Recall, that in the time-varying Linear Market Model (6.1), the time-varying systematic covariance (β_{mt}) risk can be expressed as

$$\boldsymbol{\alpha}_{1t} = \boldsymbol{\beta}_{mt}, \tag{6.4}$$

which is an extension of the result proved in section 2.1.2 in a univariate form, which can then be extended to the multivariate form here.

In this chapter we consider four special forms of (6.1) to (6.2) which make different assumptions about the data. The four alternative forms are labelled A to D and are as follows.

Model A

This model is based on the proposition that the 9 individual developed or emerging countries' stock markets are uncorrelated which is the multivariate equivalent to the univariate analysis in Chapter 5. This is achieved by defining Q and H as (9×9) diagonal matrices with their own stock market specific variances. In addition, Φ is a (9×9) diagonal matrix including an individual temporal autocorrelation parameter ϕ_i for each stock market, which allows for the different levels of temporal autocorrelation in the time-varying systematic covariance risks in the state equation (6.2) for each stock market. Thus, the parameters in this model are defined as follows.

$$\begin{aligned} H_{(9\times 9)} &= \operatorname{diag}(H_1, \ldots, H_9), \\ Q_{(9\times 9)} &= \operatorname{diag}(Q_1, \ldots, Q_9), \\ \Phi_{(9\times 9)} &= \operatorname{diag}(\phi_1, \ldots, \phi_9). \end{aligned} \tag{6.5}$$

Model B

Model B is a simplification of Model A, which retains the assumption that the individual countries' stock markets are uncorrelated. The difference between this model and Model A is that Φ is a diagonal matrix including a common temporal autocorrelation parameter ϕ for every country's stock market, instead of allowing a market specific ϕ_i as in Model A. This model thus enforces the same level of temporal autocorrelation in the time-varying systematic covariance risks for each stock market, which borrows strength in its estimation. The parameters in this model are represented by

$$\begin{aligned} H_{(9\times 9)} &= \operatorname{diag}(H_1, \ldots, H_9), \\ Q_{(9\times 9)} &= \operatorname{diag}(Q_1, \ldots, Q_9), \\ \Phi_{(9\times 9)} &= \operatorname{diag}(\phi, \ldots, \phi), \end{aligned} \tag{6.6}$$

and this model is only likely to be appropriate if the temporal autocorrelations are similar across stock markets.

Model C

Model C extends Model A by allowing for the likely correlation present between the time-varying systematic covariance risks ($\boldsymbol{\alpha}_{1t}$) of the 9 developed or emerging markets. This between stock market correlation is modelled at the level of the state equation; so the observation equation is based on the assumption of independence and hence H is a diagonal matrix with stock market specific variances as in (6.5). Different levels of temporal autocorrelation are allowed in the time-varying systematic covariance risks in the state equation, by specifying a different temporal autocorrelation parameter ϕ_i for each stock market as in (6.5). The difference between this model and Model A is that the state variance matrix Q is not diagonal, which allows for correlation between stock markets at the same time period. The specification of Q used here is given by

$$Q_{(9\times9)} = \begin{bmatrix} Q_1 & \rho\sqrt{Q_1Q_2} & \rho\sqrt{Q_1Q_3} & \cdots & \rho\sqrt{Q_1Q_9} \\ \rho\sqrt{Q_2Q_1} & Q_2 & \rho\sqrt{Q_2Q_3} & \cdots & \rho\sqrt{Q_2Q_9} \\ \rho\sqrt{Q_3Q_1} & \rho\sqrt{Q_3Q_2} & Q_3 & \cdots & \rho\sqrt{Q_3Q_9} \\ & & & \ddots & \\ \rho\sqrt{Q_9Q_1} & \rho\sqrt{Q_9Q_2} & \rho\sqrt{Q_9Q_3} & \cdots & Q_9 \end{bmatrix}, \quad (6.7)$$

which allows for a separate variance parameter for each stock market, but a constant correlation between stock markets. This constant correlation was assumed for computational simplicity and because no other factors were identified that would suggest how the correlation should be modified depending on the two stock markets in question. For example, a spatial correlation function of distance between each pair of countries could be specified, but that was considered not to be appropriate for financial data on this scale. Thus, this model takes into account two separate sources of correlation in the time-varying systematic covariance risks, the correlation over time within the same stock market (captured by $\phi_1, \ldots, \phi_9$) and the correlation in the systematic covariance risk at the same time period between stock markets (captured by ρ).

Model D

Model D extends Model B by allowing for the likely correlation between the time-varying systematic covariance risks of the individual countries' stock markets. In particular, as for Model B, H is assumed to be a diagonal matrix with stock market specific variances and Φ is a diagonal matrix including a common ϕ for each stock market, thereby enforcing the same level of temporal autocorrelation in the time-varying systematic risks for each stock market. The difference between this model and Model B is that Q allows for between stock market correlation, and is defined by (6.7) as in Model C.

Notation

Here, using the notation of the linear Gaussian state space model (3.22) and (3.23), we have

$$\begin{aligned} \boldsymbol{Y}_t &= \boldsymbol{R}_t - \boldsymbol{R}_{ft}, \\ A_t &= \left((R_{mt} - R_{ft}) \otimes I_9\right), \\ \boldsymbol{\alpha}_t &= \boldsymbol{\alpha}_{1t}, \\ \Phi &= \Phi, \end{aligned}$$

in (6.1) and (6.2). The estimation of the state vector and the remaining parameters vector $\boldsymbol{\Theta} = (\boldsymbol{\kappa}, \bar{\boldsymbol{\alpha}}_1, \Phi, Q, H)$ is as described in section 3.3. Software to implement these multivariate state space models using the Kalman Filter algorithm is not generally available, so code was written in R software as part of this PhD thesis. The algorithm used is a modified version of that described in Shumway and Stoffer (2006), and is included in Appendix A of this thesis.

6.3 Comparison of Models

6.3.1 In-sample Model Fit

This section compares the in-sample model fit performance for the four multivariate state space forms of the time-varying Linear Market Model described in section 6.2 and is applied separately to the 9 developed and 9 emerging markets. The comparisons in the model performance is achieved in terms of two different measures of error, which are the Mean Absolute Error (MAE) and the Mean Square Error (MSE), as outlined in section 3.5.1. Tables 6.2 and 6.3 display the MAE and MSE values respectively, across all 18 stock markets for each model produced by the in-sample procedure. Note that median MAE and MSE values for all four models are calculated here in addition to the mean because relatively

Table 6.2: MAE ($\times 10^2$) of in-sample model fit.

Model	A	B	C	D
Developed				
France	0.786	0.780	0.480	**0.479**
Germany	0.796	0.794	0.633	**0.630**
Italy	1.046	1.046	0.772	**0.769**
Japan	1.247	**1.246**	1.643	1.643
Norway	1.707	**1.621**	1.768	1.755
Sweden	1.178	1.191	1.163	**1.159**
Switzerland	0.941	0.933	0.843	**0.835**
UK	0.789	**0.782**	0.797	0.791
USA	**0.423**	**0.423**	0.598	0.601
Mean	0.990	0.980	0.966	**0.962**
Median	0.941	0.933	0.797	**0.791**
Emerging				
Brazil	1.671	1.591	1.573	**1.512**
Chile	1.427	1.423	1.415	**1.404**
India	2.260	1.961	1.924	**1.886**
Korea	**1.725**	1.748	1.822	1.847
Malaysia	1.126	1.139	**1.100**	1.107
Mexico	1.187	**1.181**	1.219	1.215
Poland	1.845	1.831	1.756	**1.732**
Russia	2.176	2.093	2.074	**2.039**
SouthAfrica	1.619	1.632	1.554	**1.536**
Mean	1.671	1.622	1.604	**1.586**
Median	1.671	1.632	1.573	**1.536**

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MAE.

big differences exist between the stock market MAE and MSE values.

A comparison of the four multivariate state space model formulations results display that Model C and Model D are preferable in 12 (5 developed and 7 emerging markets) out of 18 stock markets. This suggests that accounting for the correlation in the time-varying systematic covariance risks between stock markets aids in estimating stock market excess returns. In addition, the performance of Model A, which ignores the between stock market correlation and is essentially a multivariate equivalent to the univariate models described in Chapter 5, is generally worse than the other models in both developed and emerging markets. Overall, Model D (lowest MAE and MSE) improves on Model A (highest MAE and MSE) in terms of MAE (MSE) on average (median) by 15.9% (23.5%) for

Table 6.3: MSE ($\times 10^4$) of in-sample model fit.

Model	A	B	C	D
Developed				
France	1.116	1.103	**0.492**	**0.492**
Germany	1.109	1.108	0.733	**0.725**
Italy	1.949	1.949	1.164	**1.155**
Japan	2.675	**2.673**	4.673	4.674
Norway	5.207	**4.737**	5.564	5.491
Sweden	**2.404**	2.456	2.438	2.428
Switzerland	1.505	1.484	1.174	**1.152**
UK	1.170	1.159	1.142	**1.125**
USA	**0.307**	0.308	0.643	0.648
Mean	1.938	**1.886**	2.003	1.988
Median	1.505	1.484	1.164	**1.152**
Emerging				
Brazil	4.901	4.580	4.473	**4.236**
Chile	3.550	3.523	3.444	**3.422**
India	8.773	6.746	6.590	**6.440**
Korea	5.596	**5.766**	6.029	6.277
Malaysia	2.306	2.357	**2.212**	2.240
Mexico	2.404	**2.386**	2.519	2.499
Poland	6.104	6.032	5.539	**5.392**
Russia	8.910	8.266	8.058	**7.849**
SouthAfrica	4.399	4.472	4.046	**3.977**
Mean	5.216	4.903	4.768	**4.704**
Median	4.901	4.580	4.473	**4.236**

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MSE.

developed markets, and by 8.1% (13.6%) for emerging markets. This means that the relative improvement in performance of Model D compared to Model A is higher in developed markets than in emerging markets. This may be due to highly correlated developed markets, which Table 6.1 shows have greater between stock market correlations than the emerging markets. Also, it can be clearly seen that the performance of all four models in the emerging markets is worse than that for the developed markets. This may be due to the fact that the developed markets are more stable than the emerging markets, as evidenced in Chapter 5.

It is worth noting that the comparisons between Model A (including ϕ_i) and Model B (including ϕ), and Model C (including ϕ_i) and Model D (including ϕ) have been made to evaluate the consistency of the temporal autocorrelation in the time-varying systematic covariance risks for each stock market. The results of the MAE and MSE display that there is not a large difference in model fit from making the simplification that $\phi_i = \phi$ for both developed and emerging markets. To summarise, Model D which is the simplification of Model C with $\phi_i = \phi$, seems to be preferable when modelling the weekly time-varying systematic covariance (*beta*) risk for both developed and emerging markets overall, as it has the lowest MAE and MSE on average (median).

6.3.2 Out-of-sample Forecasting

A comparison of forecasting performances is now made between the same models, using an out-of-sample procedure which allows the models' predictive performance to be evaluated. A rolling window technique is considered here to undertake the out-of-sample forecasting comparison, and more theoretical details of this technique are provided in section 4.4.2. In this case, the length of the rolling window is 5 years (260 weeks approximately equal to half of the data presented), and is used to predict $\boldsymbol{\alpha}_{1t} = \boldsymbol{\beta}_{mt}$ one-week ahead (one-step ahead prediction). The length of the prediction period is 2 years (104 weeks) over the period from July 28, 2010 to July 18, 2012, which is short enough to reflect current stock market conditions. The MAE and MSE values between predicted excess returns and the actual excess returns on the stock markets are computed over these 104 values for all 18 stock markets for each model. The MAE and MSE values across the developed and emerging markets for all models are presented in Tables 6.4 and 6.5, respectively.

A comparison of the four models show that Model C and Model D are better able to predict developed and emerging markets excess returns overall, based on the median MAE and MSE values across the 9 stock markets. However, the individual stock market results suggest that the correlation in the systematic covariance risk between stock markets at the same point in time is beneficial to forecast stock market excess returns in all emerging markets, but only in 4 of the 9 developed markets. The problem here may be Japan, which exhibits relatively low correlation with the other developed markets (see in Table 6.1).

Table 6.4: MAE ($\times 10^2$) of out-of-sample forecasts.

Model	A	B	C	D
Developed				
France	0.845	0.845	**0.514**	0.520
Germany	0.764	0.774	0.595	**0.591**
Italy	1.396	1.398	1.113	**1.112**
Japan	1.165	**1.163**	1.465	1.486
Norway	1.231	**1.226**	1.456	1.442
Sweden	**1.055**	1.058	1.109	1.117
Switzerland	0.876	0.875	0.867	**0.866**
UK	**0.619**	0.620	0.735	0.733
USA	**0.358**	**0.358**	0.509	0.514
Mean	**0.923**	0.924	0.929	0.931
Median	0.876	0.875	0.867	**0.866**
Emerging				
Brazil	1.537	1.539	**1.372**	1.388
Chile	1.346	1.346	1.322	**1.319**
India	2.224	2.121	**1.900**	1.911
Korea	**1.374**	**1.374**	1.388	1.389
Malaysia	0.939	0.939	0.864	**0.859**
Mexico	0.943	0.943	**0.917**	0.919
Poland	1.532	1.548	1.440	**1.439**
Russia	1.496	1.500	**1.348**	1.374
SouthAfrica	1.398	1.410	**1.312**	1.314
Mean	1.421	1.413	**1.318**	1.324
Median	1.398	1.410	**1.348**	1.374

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MAE.

Overall, Model C (either the lowest or one of the lowest median MAE and MSE) improves on Model A (either the highest or one of the highest median MAE and MSE) in terms of MAE (MSE) on average (median) by 1% (8.4%) for developed markets and by 3.6% (7.7%) for emerging markets. Also, the MAE and MSE values for all models in the emerging markets are higher than those in developed markets, which is possibly in response to the outliers that are more common in the emerging markets than in the developed markets.

Comparisons of Model A (including ϕ_i) to Model B (including ϕ) and Model C (including ϕ_i) to Model D (including ϕ) allow us to judge the level of consistency in the temporal autocorrelation in the time-varying systematic covariance risks for each stock market in the out-of-sample procedure. Here, it can be clearly

Table 6.5: MSE ($\times 10^4$) of out-of-sample forecasts.

Model	A	B	C	D
Developed				
France	1.200	1.201	**0.527**	0.534
Germany	1.057	1.074	**0.564**	0.565
Italy	3.387	3.390	2.320	**2.315**
Japan	**2.393**	2.391	3.700	3.771
Norway	2.407	**2.388**	3.307	3.255
Sweden	**2.091**	2.099	2.326	2.352
Switzerland	1.417	1.416	**1.298**	1.328
UK	**0.667**	0.669	0.880	0.886
USA	**0.231**	**0.231**	0.414	0.421
Mean	**1.650**	1.651	1.704	1.714
Median	1.417	1.416	**1.298**	1.328
Emerging				
Brazil	3.819	3.824	**3.232**	3.282
Chile	3.393	3.393	3.200	**3.173**
India	7.053	6.568	**5.453**	5.546
Korea	3.500	3.500	3.489	**3.481**
Malaysia	1.589	1.589	1.333	**1.324**
Mexico	1.636	1.636	**1.551**	1.554
Poland	4.155	4.254	3.669	**3.657**
Russia	4.168	4.184	**3.369**	3.445
SouthAfrica	3.137	3.173	**2.797**	2.816
Mean	3.606	3.569	**3.121**	3.142
Median	3.500	3.500	**3.232**	3.282

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MSE.

seen that there is no clear differences in prediction performance from making the simplification that $\phi_i = \phi$ for the developed markets, but that there is a small reduction in performance for emerging markets. Therefore, overall, Model C seems to be preferred when forecasting the weekly time-varying systematic covariance (*beta*) risk, as it has either the lowest or one of the lowest MAE and MSE values on average (median).

6.3.3 Developed markets without Japan

This section compares the in-sample model fit and out-of-sample forecasting performance of the same models in a multivariate context for the developed markets with Japan removed. Table 6.1 shows that Japan has the lowest correlations with the other developed markets and its inclusion may be the cause of the relative poor performance of Model C and Model D (which include between stock market correlation structure) for some of the individual developed markets as well as the mean. The MAE and MSE values across the developed markets except for Japan and for all four models in both in-sample and out-of-sample procedures are presented in Tables 6.6, 6.7 and 6.8, 6.9.

Table 6.6: MAE ($\times 10^2$) of in-sample model fit without Japan.

Model	A	B	C	D
Developed				
France	0.785	0.780	0.480	**0.479**
Germany	0.796	0.795	0.633	**0.629**
Italy	1.046	1.046	0.772	**0.768**
Norway	1.707	**1.622**	1.768	1.756
Sweden	1.178	1.192	1.163	**1.159**
Switzerland	0.941	0.933	0.843	**0.835**
UK	0.789	**0.782**	0.797	0.790
USA	**0.423**	0.423	0.598	0.600
Mean	0.958	0.947	0.882	**0.877**
Median	0.869	0.864	**0.785**	0.779

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MAE.

Table 6.7: MSE ($\times 10^4$) of in-sample model fit without Japan.

Model	A	B	C	D
Developed				
France	1.116	1.104	**0.491**	0.492
Germany	1.110	1.108	0.733	**0.724**
Italy	1.949	1.949	1.164	**1.155**
Norway	5.209	**4.739**	5.568	5.500
Sweden	**2.404**	2.458	2.438	2.427
Switzerland	1.506	1.484	1.175	**1.153**
UK	1.171	1.159	1.142	**1.123**
USA	**0.307**	0.308	0.643	0.647
Mean	1.847	1.789	1.669	**1.653**
Median	1.339	1.322	1.153	**1.138**

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest MSE.

Table 6.8: MAE ($\times 10^2$) of out-of-sample forecasts without Japan.

Model	A	B	C	D
Developed				
France	0.845	0.845	**0.515**	0.520
Germany	0.765	0.773	**0.595**	0.596
Italy	1.396	1.393	**1.112**	1.121
Norway	1.231	**1.227**	1.459	1.443
Sweden	**1.054**	1.058	1.110	1.120
Switzerland	0.872	0.870	**0.868**	0.870
UK	**0.617**	**0.617**	0.736	0.734
USA	**0.358**	**0.358**	0.510	0.511
Mean	0.892	0.893	**0.863**	0.864
Median	0.859	0.858	**0.802**	**0.802**

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest *MAE*.

Table 6.9: MSE ($\times 10^4$) of out-of-sample forecasts without Japan.

Model	A	B	C	D
Developed				
France	1.202	1.201	**0.530**	0.536
Germany	1.057	1.070	**0.566**	0.568
Italy	3.382	3.372	**2.321**	2.346
Norway	2.410	**2.393**	3.321	3.256
Sweden	**2.092**	2.101	2.328	2.354
Switzerland	1.409	1.406	**1.305**	1.338
UK	**0.663**	0.663	0.884	0.888
USA	**0.231**	**0.231**	0.415	0.416
Mean	1.556	1.555	**1.459**	1.463
Median	1.306	1.304	**1.095**	1.113

Notes: **Bold** displays the best market pricing model for each stock market in terms of the lowest *MSE*.

A comparison of the four models in both the in-sample and out-of-sample procedures show that Model C and Model D which incorporate between stock market correlation seem to be preferable for modelling and forecasting the weekly time-varying systematic risks for the developed markets based on both the mean and median MAE and MSE values, which was not the case when Japan was included. In addition, a comparison of the developed markets with Japan (see in Tables 6.2, 6.3, 6.4 and 6.5) and without Japan (see in Tables 6.6, 6.7, 6.8 and 6.9) show that the mean and median MAE and MSE values have decreased now Japan has been removed in both the in-sample and out-of-sample procedures. However, individually the MAE and MSE values for each developed market without Japan display the fact that there are no clear differences compared with including Japan. This may be because Japan has no clear impact when estimating models parameters in a multivariate context, as it was only one of the 9 stock market data sets modelled. This will be discussed in next section 6.4. Thus, the reduction in the average MAE and MSE values appears to be due to simply removing the series with the largest MAE and MSE values, rather than improving the fit of the models in the remaining markets.

6.4 Best Forecasting Model

This section focuses on the best forecasting model for developed and emerging markets, which is important for making profitable investment decisions for stock market investors in the future. Model C exhibited the best prediction performance overall and it is examined in greater detail here. Tables 6.10 and 6.11 represent the hyperparameter estimates of Model C using KFMR, which is defined in section 6.2. The table shows developed markets with Japan (top panel) and without Japan (middle panel), as well as emerging markets (bottom panel).

The average estimated $\hat{H}_i$ and $\hat{Q}_i$ values for emerging markets are generally higher than those of developed markets, which may be due to emerging markets' being more volatile. This may be due to the variable involved in capturing financial and economic integration, such as the trade volume of stock markets. For example, the trade volume of developed markets is greater than that of emerging markets. This suggests that emerging markets are more vulnerable to financial speculation (e.g. the buying and selling of stocks, bonds, currency etc.) than developed markets; thus, the fluctuation in emerging markets is anticipated to be larger than that in developed markets. Also, the average estimated values of $\hat{Q}_i$ are higher than those of $\hat{H}_i$, meaning that the state variance captures the volatility of the stock market excess returns more than the observation variance. It is worth noting that the estimated $\hat{Q}_i$ for the USA is 0 (3 decimals), meaning that systematic covariance risk is constant over time.

Table 6.10: Model C parameter estimates (standard errors) via KFMR for the developed markets.

Market	$\hat{Q}_i \times 100$	$\hat{H}_i \times 100$	$\hat{\rho}$	$\hat{\phi}_i$	$\hat{\kappa}_i$	$\hat{\bar{\alpha}}_{1i}$
Developed						
France	43.651	0.007	0.892	0.154	0.001	1.298
	(6.775)	*(0.001)*	*(0.179)*	*(0.015)*	*(0.000)*	*(0.056)*
Germany	53.628	0.010	0.892	0.129	0.001	1.343
	(7.843)	*(0.001)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.064)*
Italy	71.253	0.016	0.892	0.110	0.000	1.228
	(11.613)	*(0.001)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.068)*
Japan	0.008	0.047	0.892	0.000	0.000	0.682
	(0.062)	*(0.003)*	*(0.179)*	*(0.000)*	*(0.000)*	*(0.024)*
Norway	69.323	0.062	0.892	0.024	0.003	1.386
	(15.385)	*(0.005)*	*(0.179)*	*(0.008)*	*(0.000)*	*(0.092)*
Sweden	41.231	0.028	0.892	0.077	0.002	1.415
	(8.216)	*(0.002)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.069)*
Switzerland	25.783	0.014	0.892	0.282	0.001	0.904
	(5.485)	*(0.001)*	*(0.179)*	*(0.037)*	*(0.000)*	*(0.038)*
UK	13.957	0.013	0.892	0.191	0.001	1.075
	(3.298)	*(0.001)*	*(0.179)*	*(0.030)*	*(0.000)*	*(0.035)*
USA	0.000	0.006	0.892	0.988	0.000	0.913
	(0.000)	*(0.000)*	*(0.179)*	*(0.552)*	*(0.000)*	*(0.015)*
Average	**35.426**	**0.023**	**0.892**	**0.217**	**0.001**	**1.138**
Without Japan						
France	43.684	0.007	0.892	0.153	0.001	1.298
	(6.776)	*(0.001)*	*(0.179)*	*(0.015)*	*(0.000)*	*(0.056)*
Germany	53.678	0.010	0.892	0.129	0.001	1.343
	(7.846)	*(0.001)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.064)*
Italy	71.312	0.016	0.892	0.109	0.000	1.229
	(11.618)	*(0.001)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.068)*
Norway	69.277	0.062	0.892	0.024	0.003	1.386
	(15.389)	*(0.005)*	*(0.179)*	*(0.008)*	*(0.000)*	*(0.092)*
Sweden	41.247	0.028	0.892	0.077	0.002	1.415
	(8.218)	*(0.002)*	*(0.179)*	*(0.014)*	*(0.000)*	*(0.069)*
Switzerland	25.784	0.014	0.892	0.281	0.001	0.904
	(5.486)	*(0.001)*	*(0.179)*	*(0.037)*	*(0.000)*	*(0.038)*
UK	13.982	0.013	0.892	0.190	0.001	1.075
	(3.300)	*(0.001)*	*(0.179)*	*(0.030)*	*(0.000)*	*(0.035)*
USA	0.000	0.006	0.892	0.988	0.000	0.913
	(0.000)	*(0.000)*	*(0.179)*	*(0.552)*	*(0.000)*	*(0.015)*
Average	**39.871**	**0.020**	**0.892**	**0.244**	**0.001**	**1.195**

Notes: *Italic* numbers in parentheses denote the standard errors of Model C parameter estimates via KFMR for the developed markets.

Table 6.11: Model C parameter estimates (standard errors) via KFMR for the emerging markets.

Market	$\hat{Q}_i \times 100$	$\hat{H}_i \times 100$	$\hat{\rho}$	$\hat{\phi}_i$	$\hat{\kappa}_i$	$\hat{\bar{\alpha}}_{1i}$
Emerging						
Brazil	106.051	0.060	0.635	0.326	0.005	1.483
	(17.981)	*(0.005)*	*(0.051)*	*(0.040)*	*(0.000)*	*(0.140)*
Chile	35.515	0.041	0.635	0.312	0.003	0.884
	(7.620)	*(0.003)*	*(0.051)*	*(0.061)*	*(0.000)*	*(0.054)*
India	62.870	0.078	0.635	0.405	0.003	1.102
	(19.029)	*(0.007)*	*(0.051)*	*(0.097)*	*(0.000)*	*(0.098)*
Korea	91.725	0.075	0.635	0.000	0.002	1.223
	(16.472)	*(0.006)*	*(0.051)*	*(0.000)*	*(0.000)*	*(0.095)*
Malaysia	24.665	0.026	0.635	0.000	0.002	0.597
	(5.572)	*(0.002)*	*(0.051)*	*(0.000)*	*(0.000)*	*(0.026)*
Mexico	38.766	0.031	0.635	0.146	0.003	1.189
	(6.892)	*(0.003)*	*(0.051)*	*(0.025)*	*(0.000)*	*(0.064)*
Poland	108.420	0.071	0.635	0.118	0.003	1.427
	(19.554)	*(0.006)*	*(0.051)*	*(0.020)*	*(0.000)*	*(0.121)*
Russia	112.096	0.100	0.635	0.302	0.003	1.416
	(21.917)	*(0.008)*	*(0.051)*	*(0.047)*	*(0.000)*	*(0.146)*
SouthAfrica	54.385	0.050	0.635	0.202	0.003	1.253
	(10.858)	*(0.004)*	*(0.051)*	*(0.031)*	*(0.000)*	*(0.084)*
Average	**70.499**	**0.059**	**0.635**	**0.201**	**0.003**	**1.175**

Notes: *Italic* numbers in parentheses denote the standard errors of Model C parameter estimates via KFMR for the emerging markets.

The average temporal autocorrelation in the time-varying systematic covariance risk is similar for the developed markets and emerging markets, with average values of 0.217 and 0.201 respectively. These values are much closer to 0 than 1, suggesting that the time-varying systematic covariance risks change rapidly due to the low autocorrelation. The exception to this is the USA, which has $\hat{\phi}_i = 0.988$. This, taken with $\hat{Q}_i = 0.000$, suggests that its systematic covariance risk is constant over time. The common correlation in the time-varying systematic covariance risk between stock markets (captured by $\hat{\rho}$) in developed markets (0.892 with a standard error of 0.179) is higher than that of the emerging markets (0.635 with a standard error of 0.051). This may be explained by volatility, which is lower in developed markets hence resulting in higher between stock market correlations.

The average regression intercept, $\hat{\kappa}_i$ (proxy for the unexpected risk) is close to

0 in developed and emerging markets, which likely to be a consequence of the risk-free rate's (R_{ft}) being subtracted before estimation (see Campbell et al. (1997)). The mean of the time-varying systematic covariance risk $\hat{\alpha}_{1it}$ for developed and emerging markets is positive, and is greater than 1, indicating that the stock market is more volatile than the MSCI World market portfolio.

It is worth noting that a sensitivity analysis for developed markets including Japan (top panel) and with Japan removed (middle panel) shows that on average that estimate for $\hat{Q}_i$, $\bar{\hat{\alpha}}_{1i}$ and $\hat{\phi}_i$ increase and that the average for $\hat{H}_i$ decreases, while common $\hat{\rho}$ and average $\hat{\kappa}_i$ are the same when Japan is included. However, the individual parameter estimates in each stock market are the same when Japan is removed, meaning that the individual parameter estimates in developed markets including Japan are likely to be consistent with or without it. This suggests a certain robustness to the choice of countries included in the model, even for the models that account for between stock market correlations.

Table 6.12: Univariate diagnostic test statistics for Model C via KFMR.

Market	JB	$Het(174)$	$LB(23)$	Market	JB	$Het(174)$	$LB(23)$
Developed				*Emerging*			
France	3807.92*	0.54	27.65	Brazil	73.74*	0.45	16.38
	(0.000)	*(0.999)*	*(0.090)*		*(0.000)*	*(0.999)*	*(0.632)*
Germany	589.46*	0.50	25.85	Chile	76.91*	0.67	26.34
	(0.000)	*(0.999)*	*(0.134)*		*(0.000)*	*(0.996)*	*(0.121)*
Italy	732.54*	1.78*	13.44	India	133.07*	0.82	22.37
	(0.000)	*(0.000)*	*(0.815)*		*(0.000)*	*(0.905)*	*(0.266)*
Japan	65.93*	0.81	28.34	Korea	441.09*	0.66	24.31
	(0.000)	*(0.917)*	*(0.077)*		*(0.000)*	*(0.997)*	*(0.185)*
Norway	73.03*	0.57	28.98	Malaysia	276.25*	0.59	22.81
	(0.000)	*(0.999)*	*(0.066)*		*(0.000)*	*(0.999)*	*(0.246)*
Sweden	173.79*	0.94	28.46	Mexico	28.72*	0.56	34.10*
	(0.000)	*(0.658)*	*(0.075)*		*(0.000)*	*(0.999)*	*(0.018)*
Swit.	49.20*	0.79	23.84	Poland	90.26*	0.78	42.21*
	(0.000)	*(0.940)*	*(0.202)*		*(0.000)*	*(0.949)*	*(0.002)*
UK	363.44*	0.64	35.76*	Russia	368.70*	0.33	26.96
	(0.000)	*(0.998)*	*(0.011)*		*(0.000)*	*(0.999)*	*(0.106)*
USA	94.60*	0.98	55.84*	S.Africa	59.36*	0.73	33.42*
	(0.000)	*(0.553)*	*(0.000)*		*(0.000)*	*(0.980)*	*(0.021)*

Notes: The abbreviations of countries are: Swit: Switzerland and S.Africa: South Africa.
JB is the Jarque-Bera test statistic for the null hypothesis of normally distributed standardised residuals. *JB* follows χ^2 with 2 degrees of freedom so the critical value at the 5% level is 5.99.
LB(23) is the Ljung-Box test statistic for the null hypothesis of no autocorrelation in the standardised residuals up to order $\sqrt{522} \approx 23$. *LB(23)* statistic follows χ^2 with 23-$(m-1)$ degrees of freedom where m is the total number of estimated parameters.
Het(174) is the test statistic for the null hypothesis of no heteroskedasticity in the standardised residuals up to order $522/3 = 174$. *Het(174)* statistic follows $F_{(174,174)}$ distribution so the critical value at the 5% level is 1.28. * means the appropriate null hypothesis is rejected at the 5% significance level.

Table 6.12 presents the univariate diagnostic test statistics for the residuals from Model C via KFMR outlined in section 3.5.2.1. According to the Jarque-Bera (JB) test, the residuals are not normally distributed at the 5% significance level for all 18 global markets, implying that Model C is poor in terms of non-normal errors. According to the H ($Het(174)$) test, the null hypothesis of no heteroskedasticity cannot be rejected for all 18 global markets except for Italy at the 5% significance level, implying that Model C is adequate in terms of no heteroskedasticity. According to the Ljung-Box ($LB(23)$) test, the null hypothesis of no autocorrelation cannot be rejected at the 5% significance level for 13 out of the 18 global markets, meaning that Model C is not adequate in terms of no autocorrelation.

Table 6.13: Multivariate diagnostic test statistics for Model C via KFMR.

Market	MJB	$MHet(1566)$	$MLB(23)$
Developed	325.47*	0.96	2075.49*
	(0.000)	*(0.790)*	*(0.000)*
Emerging	135.68*	1.08	1907.85
	(0.000)	*(0.064)*	*(0.089)*

Notes: *MJB* is the multivariate Jarque-Bera test statistic for the null hypothesis of multivariate normally distributed standardised residuals. *MJB* follows χ^2 with 2×9 degrees of freedom so the critical value at the 5% level is 28.87.
MLB(23) is the multivariate Ljung-Box test statistic for the null hypothesis of no multivariate autocorrelation in the standardised residuals up to order $\sqrt{522} \approx 23$. *MLB(23)* statistic follows χ^2 with $(9^2 \times 23)$-m degrees of freedom where m is the total number of estimated parameters.
MHet(1566) is the test statistic for the null hypothesis of no multivariate heteroskedasticity in the standardised residuals up to order $9 \times 522/3 = 1566$. *MHet(1566)* statistic follows $F_{(1566,1566)}$ distribution so the critical value at the 5% level is 1.09. * means the appropriate null hypothesis is rejected at the 5% significance level.

Table 6.13 presents the multivariate diagnostic test statistics for the residuals from Model C via KFMR outlined in section 3.5.2.2. According to the multivariate Jarque-Bera (MJB) test, the residuals are not multivariate normally distributed at the 5% significance level for both the developed and emerging markets, implying that Model C is poor in terms of multivariate non-normal errors. According to the multivariate H ($MHet(1566)$) test, the null hypothesis of no multivariate heteroskedasticity cannot be rejected at the 5% significance level for both the developed and emerging markets, implying that Model C is adequate in terms of multivariate homoskedasticity. According to the multivariate Ljung-Box ($MLB(23)$) test, the null hypothesis of no multivariate autocorrelation is rejected at the 5% significance level for the developed markets, but not for the emerging markets, meaning that Model C is not adequate in terms of no multivariate autocorrelation for the developed markets.

6.5 Extension of Best Forecasting Model

Previously, we found that Model C exhibited the best prediction performance overall and it was examined in greater detail in section 6.4. Here, we suggest a new model called Model E which extends Model C by allowing for the likely different levels of correlation present between the time-varying systematic covariance risks of the 9 developed or emerging markets (captured by ρ_{ij}, $(i, j = 1, \ldots, 9)$). In particular, for Model C, H is assumed to be a diagonal matrix with stock market specific variances, and Φ is a diagonal matrix including the different levels of temporal autocorrelation in the time-varying systematic covariance risks for each stock market (captured by $\phi_1, \ldots, \phi_9$). The difference between this model and Model C is that the state variance matrix Q (equation (6.7)) allows for the different levels of correlation in the time-varying systematic covariance risks during the same time period between stock markets, and is defined as

$$Q_{(9\times 9)} = \begin{bmatrix} Q_1 & \rho_{12}\sqrt{Q_1Q_2} & \rho_{13}\sqrt{Q_1Q_3} & \cdots & \rho_{19}\sqrt{Q_1Q_9} \\ \rho_{12}\sqrt{Q_1Q_2} & Q_2 & \rho_{23}\sqrt{Q_2Q_3} & \cdots & \rho_{29}\sqrt{Q_2Q_9} \\ \rho_{13}\sqrt{Q_1Q_3} & \rho_{23}\sqrt{Q_2Q_3} & Q_3 & \cdots & \rho_{39}\sqrt{Q_3Q_9} \\ & & & \ddots & \\ \rho_{19}\sqrt{Q_1Q_9} & \rho_{29}\sqrt{Q_2Q_9} & \rho_{39}\sqrt{Q_3Q_9} & \cdots & Q_9 \end{bmatrix}. \quad (6.8)$$

Here we enforce symmetry on Q so that $\rho_{ij}=\rho_{ji}$ ($i, j = 1, \ldots, 9$).

Note that the *optim* package in R is used throughout this thesis while estimating the unknown parameters of the state space models. Here, we again focused on the estimation of the unknown parameters of Model E through the loglikelihood function which can be computed and maximized by a Newton-Raphson numerical search algorithm via the *optim* package (see details in section 3.3). However, *optim* package failed to converge in the optimization step while maximizing the loglikelihood function of Model E. Petris (2014) suggests that this may be due to the large number of unknown parameters of the multivariate state space model via the Kalman Filter. Hence, Model E is not discussed further in this thesis.

6.6 Conclusion

The focus of this chapter was on modelling and forecasting time-varying systematic covariance (*beta*) risks, using a multivariate state space form of the time-varying Linear Market Model (consistent with conditional Two-Moment CAPM) using a KFMR. In particular, the aim was to see if the multivariate context which accounted for between stock market correlations outperforms a univariate approach, using both in-sample and out-of-sample procedures. The modelling and forecasting abilities of four multivariate state space forms of the time-varying Linear Market Model as described in section 6.2 were applied separately to the 9 developed and 9 emerging markets. In addition, the developed markets were also modelled without Japan in a sensitivity analysis, because it had the lowest correlations within the developed markets (see in Table 6.1), and may have affected model performance.

The performance of these models, when using the in-sample and out-of-sample procedures, were evaluated using MAE and MSE. Overall, Model A, which ignores the correlation structures and applies a multivariate context equivalent to the univariate context in Chapter 5, is generally worse than the other models in both developed (with and without Japan) and emerging markets. Model B simplifies the temporal autocorrelation in the time-varying systematic covariance risks; so that $\phi_i = \phi$ in for both developed and emerging markets. This simplification does not greatly reduce the performance compared to Model A in both the in-sample and out-of-sample procedures, which may suggest that either the stock markets exhibit common temporal autocorrelation (ϕ), or that the use of 9 stock markets data improves the estimation of ϕ_i. However, the key finding from this chapter is that incorporating between stock market correlations into the modelling, as in Model C and Model D, improves both the in-sample modelling and out-of-sample forecasting abilities of the time-varying Linear Market Model. This latter result will be of most interest to researchers and stock market investors, who are concerned with forecasting future stock market movements to make profitable investment in stock markets. This study is one of the few in the financial literature to undertake a multivariate stock market approach which utilizes between stock market correlation, as the majority of the financial models of this type are univariate and treat each market individually. The presence of such between stock market correlation is not surprising, especially in the developed economies which are becoming increasingly financially integrated. For example, the increased financial integration of the Eurozone countries using the same currency may result in increased trade flow within these countries. The relative improvements in the performance of Models C and D compared to Model A (univariate context) in developed markets was generally higher than in emerging markets in the in-sample procedure. This may be explained by stock market correlations, which are higher in developed markets, as displayed in Table 6.1 hence resulting in greater improvements to the performance of Models C and D. The differences in model performance between Model C and Model D were generally slight and not universally consistent, which suggests that incorporating single (ϕ) or market specific (ϕ_i) temporal autocorrelation parameters does not make a big impact on model performance. This is likely to be because Table 6.10 shows that when Model C is fitted (market specific (ϕ_i) temporal autocorrelations) the estimated temporal autocorrelation parameters are all similar and close to 0.2, with the exception of the USA which is estimated as 0.988.

The sensitivity analysis for developed markets with and without Japan highlights that the country specific estimation performance of the majority of markets is likely to be consistent in the presence of a single outlying country. This suggests a certain robustness of prediction performance to the choice of countries included in the model, even for the models that account for between stock market correlations. However, the estimation for the outlying country itself is relatively poor compared with assuming all countries are independent. This suggests that if multivariate modelling is to be undertaken, then the choice of a sensible set of countries to model is crucial to ensure good prediction performance for all (and not just the majority) countries' markets. Thus if this PhD was longer than it would be interesting to attempt to group countries together based on similar financial markets, which would then aid in the multivariate modelling considered here.

Note that residual diagnostics are discussed in section 6.4. These assumptions include that the residuals are multivariate normally distributed, independent (no multivariate autocorrelation) and have constant variance (multivariate homoskedasticity) for the multivariate context. When these assumptions are violated, the performance of the multivariate state space model can be affected, so possible extensions of the multivariate state space model are discussed as follows.

The first assumption is that the residuals are multivariate Gaussian distributed. This assumption is violated, which is a consequence of the asymmetry and the heavy tails of the multivariate case. The multivariate Gaussian distribution can be replaced by multivariate heavy-tailed distributions such as the multivariate t distribution, or a mixture of multivariate normals, or by a multivariate asymmetric distribution, such as a multivariate skewed-t distribution.

The second assumption is that the variance of residuals is assumed to be constant (multivariate homoskedasticity) for the multivariate context. When this assumption is violated, a multivariate stochastic volatility model allows a combined GARCH-type and state space model to capture the time-varying variance of residuals for the multivariate context. The various extensions of multivariate stochastic volatility models exist not only with multivariate Gaussian residuals but also with multivariate non-Gaussian residuals.

The final assumption is that the residuals are independent (no multivariate autocorrelation) for the multivariate context. When this assumption is invalid, we would suggest an approach which allows the intercept vector to vary over time via a random walk model within the Kalman Filter algorithm, in addition to a slope parameter vector $\boldsymbol{\alpha}_{1t} = \boldsymbol{\beta}_{mt}$.

Chapter 7

Conclusion and Further Work

7.1 Summary

This thesis focuses on financial and statistical modelling of returns in financial stock markets, using the concept of the asset pricing model in finance. This model is important for making investment decisions such as portfolio choice, and determining market risk for investors and researchers.

The best known and most widely used asset pricing model in the finance literature is the Two-Moment Capital Asset Pricing Model (CAPM) (consistent with the Linear Market Model) developed by Sharpe-Lintner-Mossin in the 1960s and is the benchmark model of this thesis. This model implies a linear relationship between the expected return on a financial asset and that over the whole market in which the asset is traded. The slope coefficient is called the systematic covariance (*beta*) risk, which is assumed to be constant over time. The literature (e.g. Kraus and Litzenberger (1976), Fang and Lai (1997), Hwang and Satchell (1999), Mergner and Bulla (2008) and Choudhry and Wu (2009)) however, shows that this model may be misleading and insufficient as a tool for characterising returns in financial time series data, which is possibly a consequence of the non-linearity in the relationship between returns on asset and whole market. The inadequacies of the Two-Moment CAPM motivated me to validate and extend the benchmark Linear Market Model in this thesis. In doing this, the following three research aims were addressed:

1. Evaluate the effectiveness of the widely used Two-Moment CAPM and its Data Generating Process (DGP) equivalent, the Linear Market Model, for modelling and forecasting returns in financial time series data. In doing this, I have

    (a) Assessed the extent to which the linearity assumptions underpinning the Linear Market Model are appropriate, or whether increasing its flexibility by adding higher order moments, for the purpose of allowing for skewness and kurtosis, or non-linearities could improve the fit of the model to the stock market returns data.

    (b) Assessed whether the linearity assumption of the Linear Market Model is realistic locally rather than globally, which is achieved by comparing its modelling and forecasting performance against a Linear Market Model with a time-varying systematic covariance (*beta*) risk parameter estimated using state space models.

    (c) Assessed whether the local linearity assumption of the time-varying Linear Market Model is appropriate by comparing its performance to the time-varying Quadratic Market Model and the time-varying Cubic Market Model which incorporate time-varying systematic skewness (*co-skewness*) and time-varying systematic kurtosis (*co-kurtosis*) risk parameters.

    (d) Assessed whether the constant variance assumption underpinning the Linear Market Model is appropriate by comparing its performance to that of GARCH-type models.

2. Investigate whether a multivariate approach to modelling and forecasting systematic covariance (*beta*) risk, which allows for between stock market correlations, outperforms a univariate, one stock market at a time, approach. Due to the results in earlier chapters, this investigation was undertaken in the context of allowing time-varying systematic covariance risk parameter using multivariate state space models.

3. Evaluate the ability of a variety of statistical modelling techniques to forecast recent returns in financial data across a variety of stock market conditions such as before and after the October 2008 financial crisis, including not only developed markets (which typically exhibit stable behaviour) but also emerging markets (which are more volatile and prone to large fluctuations and instabilities). In particular, in order to assess the latter, a case study on Turkey's industry sector portfolios was undertaken.

This thesis was organised as follows. The theoretical details of the financial and statistical methodologies used were provided in Chapter 2 and Chapter 3, respectively. The empirical work for investigating the research aims was presented in Chapter 4 and Chapter 5 for the univariate context, and in Chapter 6 for the multivariate context. Finally, this chapter reviews the main conclusions and key themes which affect the direction of continuing and future research connected with this thesis's findings.

Chapter 4 compared the commonly used benchmark model, the Linear Market Model, with time-varying Linear Market Models for 19 Turkish industry sector portfolios, using both in-sample and out-of-sample procedures. Two time-varying Linear Market Model specifications were considered: GARCH-type models, which allow for non-constant variance in stock market returns, and state space models, which allow for the systematic covariance risk to change linearly over time. The main conclusions from this chapter are that, while GARCH-type models do outperform the Linear Market Model, particularly in terms of their forecasting performance, the differences are not large. The state space models, however, perform consistently better than both the Linear Market Model and GARCH-type models, in terms of both modelling (in-sample MAE and MSE) and forecasting (out-of-sample MAE and MSE) performance. In particular, the mean reverting (KFMR) formulation of the state space model is superior to the random walk (KFRW) and random coefficients (KFRC) specifications, due both to its improved modelling and forecasting performance and to its increased flexibility.

Chapter 5 extended the Linear Market Model to allow for non-linearity in six main ways. The first two were polynomial extensions of the Linear Market Model as outlined in Chapter 2, namely the Quadratic Market (allowing for systematic covariance and systematic skewness) and Cubic Market Models (allowing for systematic covariance and systematic skewness, as well as systematic kurtosis). The third approach relaxed the rigid shapes enforced by polynomial models by using a generalised additive model (GAM). The next approach is the time-varying Linear Market Model compared was a state space model as used in Chapter 4. The last two approaches are the time-varying versions of polynomial extensions, namely, the time-varying Quadratic Market Model (allowing for time-varying systematic covariance and time-varying systematic skewness) and the time-varying Cubic Market Model (allowing for time-varying systematic covariance, time-varying systematic skewness and time-varying systematic kurtosis). The models were applied to the 18 global markets, including the 9 developed and 9 emerging markets, during the three different time periods: the entire period from July 2002 to July 2012, from July 2002 to before the October 2008 financial crisis, and from after the October 2008 financial crisis to July 2012. The main conclusions from this chapter are that the proposed time-varying Linear Market Model via KFMR yields the best modelling of the 18 global market returns during these three different time periods.

During the entire period from July 2002 to July 2012, the proposed time-varying Linear Market Model via KFMR improves on the time-varying Quadratic Market Model via KFMR in terms of *Adjusted* R^2 by on average 0.3% for the developed and 0.6% for the emerging markets while the proposed time-varying Linear Market Model improves on the time-varying Cubic Market Model via KFMR in terms of *Adjusted* R^2 by on average 0.5% for the developed markets and 1% for the emerging markets. Furthermore, it can be seen that the proposed time-varying Linear Market Model via KFMR improves on the Linear Market Model in terms of *Adjusted* R^2 on average by 10.4% for the developed markets and 25.3% for the emerging markets. However, the GAM outperformed the Higher order DGPs in all 18 global markets while improving on the Linear Market Model in terms of *Adjusted* R^2 on average by only 1% for the developed markets and 3.9% for the emerging markets. In addition, the Cubic Market Model provides a slight improvement on the Linear Market Model, while the modelling performance of the Quadratic Market Model is worse than that of the Linear Market Model in some stock markets such as Germany, India and Malaysia during this period.

During the periods before and after the October 2008 financial crisis, the proposed time-varying Linear Market Model provides a much better performance than any of the other models for both the developed and emerging markets, with the exception of the emerging markets during the period after the October 2008 financial crisis, where the time-varying Quadratic Market Model performs better in terms of average *Adjusted* R^2; however, the improvements in model fit to the proposed time-varying Linear Market Model are not substantial. For example, the time-varying Quadratic Market Model improves on the proposed time-varying Linear Market Model in terms of *Adjusted* R^2 by on average 0.3% for the emerging markets. It can be seen that the proposed time-varying Linear Market Model improves on the Linear Market Model in terms of *Adjusted* R^2 by on average 9.6% for the developed markets and 24.2% for the emerging markets during the period before the October 2008 financial crisis, whilst making improvements in model fit from the Linear Market Model to the proposed time-varying Linear Market Model in terms of *Adjusted* R^2 by on average 7.7% for the developed markets and 14.2% for the emerging markets during the period after the October 2008 financial crisis. In addition, the GAM outperformed the Higher order DGPs in terms of average *Adjusted* R^2 for all 18 global markets, while improving on the Linear Market Model in terms of *Adjusted* R^2. Improvements were on average 2.1% for the developed markets and 5.3% for the emerging markets in the pre-crisis period; while in the post-crisis period, improvements were on average 0.4% for the developed markets and 0.6% for the emerging markets.

The previous empirical investigations in Chapter 4 and Chapter 5 focused the evaluation in a univariate context. In contrast, Chapter 6 compares the performance of four possible multivariate state space forms of the time-varying Linear Market Model using a state space model via KFMR, again in 18 global markets in both the in-sample and out-of-sample procedures. In addition, the developed markets were also modelled without Japan in a sensitivity analysis, because it had the lowest correlations within the developed markets and may have affected the model's performance. The main conclusions from this chapter are that Models C and D, which incorporate between stock market correlations into the modelling, improve the quality of the modelling (in-sample median MAE and MSE) and the forecasting (out-of-sample median MAE and MSE) abilities of the time-varying Linear Market Model in both developed and emerging markets. In addition, the relative improvement in the in-sample performance of Models C and D, when compared to the univariate context, was generally greater in the developed markets than in the emerging markets, which is likely to be because these markets exhibit larger between market correlations and are thus more suited to a multivariate approach. It is worth noting that the differences between the performance of Model C and Model D were generally slight and not universally consistent, suggesting that incorporating a single (ϕ) or market specific (ϕ_i) temporal autocorrelation does not have a big impact on model performance. It is worth noting that the out-of-sample forecasting results were generally supportive by Model C, which is of interest to stock market investors and those researchers who are concerned with forecasting future stock market movements and making profitable investments in the stock market. In addition, the results of the sensitivity analysis for developed markets, either with or without Japan, display the fact that country specific estimation performance for the majority of markets are likely to be consistent in the presence of a single outlying country. However, the estimation for the outlying country itself was relatively poor when contrasted with the results obtained on the assumption that all countries are independent. Hence, it would be interesting to attempt to group countries together based on similar financial stock markets, thereby assisting the multivariate modelling considered here.

7.2 Key Themes

7.2.1 Inappropriateness of the Linear Market Model

The Linear Market Model, which was the benchmark model in this thesis and in the literature, performs relatively well in developed markets but not so well in emerging markets. Also, I have found that adding in higher order moments (or GAM) does not lead to large improvements over the Linear Market Model, in contrast with results from the existing literature (e.g. Fang and Lai (1997) and Hwang and Satchell (1999)) which uses them. This is may be due to the fact that the financial data used here are more stable with fewer outliers than those used in the existing literature.

Similarly, allowing for non-constant variance via GARCH-type models while improving the forecasting performance of the Linear Market Model does not do so to a large extent. However, allowing for local linear behaviour in the systematic covariance (*beta*) risk in the time-varying Linear Market Model via state space models does dramatically improve the modelling and forecasting performance when compared with the Linear Market Model. Similarly, the time-varying versions of polynomial extensions, namely, the time-varying Quadratic Market Model and the time-varying Cubic Market Model have also led to large number of improvements over the Linear Market Model, but the modelling performance of these models was generally worse than the time-varying Linear Market Model. This suggests that the relationship between returns on asset and the whole market is locally linear rather than quadratic or cubic (or any other non-linear shapes), but that the slope of the linear relationship varies over time.

7.2.2 Multivariate modelling improves univariate market modelling

Financial stock market data are correlated between stock markets. A univariate modelling approach for each stock market that ignores this correlation is relatively poor. In contrast, undertaking a multivariate modelling approach which allows for this between stock market correlation improves the modelling and forecasting performance of state space models, because it borrows strength in the estimation across similar stock markets. Higher correlations were observed between developed markets compared to emerging markets, making the multivariate approach most beneficial in this context. However, as demonstrated by the Japan stock market (which had the lowest correlations within developed markets), the choice of the markets included will affect the accuracy of the forecasting performance for an individual stock market. This between stock market correlation suggests that the use of univariate approaches in the majority of the existing financial literature is non-optimal, and that multivariate models should become the norm.

7.2.3 State of the world's stock markets

In the developed world, their financial stock markets are similar and move largely in unison, due to their increased financial integration such as Eurozone countries using the same currency, etc. This backs up the previous theme that a multivariate context is beneficial in modelling such data. Emerging markets are less correlated and more volatile; thus, in general, modelling and forecasting their values can be done with less accuracy than for developed markets. With the Turkish industry sector portfolio data the main features of these data are the positive mean, relatively high volatility, asymmetry (left-skew and right-skew) and leptokurtosis (fat tails). These findings match the most common features of emerging markets found by previous authors. This justifies the consideration of the conditional Two-Moment CAPM model for capturing time-varying variance and covariance, rather than simple the Two-Moment CAPM for each industry sector portfolio.

7.3 Further Work

There are still many interesting ideas open for further research. One idea for future work is that these financial models and statistical modelling techniques be applied to different financial stock markets, to different sample time periods (such as during a financial crisis), and to different data frequencies (such as daily, monthly, etc.). This would then further our understanding about how the modelling and forecasting performance of these models are affected by stock markets, sample time periods and data frequency, particularly the presence of temporal autocorrelation.

The poor performance of GARCH-type models which was observed may be due to the assumption of a constant correlation across time. This could be relaxed by allowing it to vary smoothly over time. Thus another possible extension could be made by considering GARCH-type models with a dynamic conditional correlation, i.e. one that varies over time.

Another possible extension could be made by using a stochastic volatility model allowing a combined GARCH-type and state space model. It would be interesting to see if the modelling and forecasting performance is affected by a combined GARCH-type and state space model for both the univariate and multivariate contexts.

Another possible extension could be made by extending the multivariate state space models considered here with time-varying Higher order DGPs (i.e. the multivariate time-varying Quadratic Market and time-varying Cubic Market Models). It would be interesting to see if the modelling and forecasting performance in the multivariate contexts is affected by incorporating Higher order DGPs.

The poor performance of all models observed in the financial literature may possibly be a consequence of the non-normally distributed returns and non-linear relationship between asset and market portfolio returns. In this thesis, we discussed how to deal with the non-linearity which exits between asset and market portfolio returns for both the univariate and multivariate contexts. Thus, it is worth considering further possible extensions which would deal with the non-normally distributed returns for both the univariate and multivariate contexts. The non-normality case is a consequence of asymmetry and heavy tails for both the univariate and multivariate contexts. Therefore, the normal distribution can be replaced by heavy-tailed distributions such as the t distribution, or a mixture of normals or a general residual distribution, or by an asymmetric distribution, such as a skewed-t distribution for the univariate context. In addition, the multivariate normal distribution can be replaced by multivariate heavy-tailed distributions such as the multivariate t distribution, or a mixture of multivariate normals, or by a multivariate asymmetric distribution, such as a multivariate skewed-t distribution for the multivariate context. Indeed, it would be interesting to observe whether these extensions would improve the modelling and forecasting performance of all models.

Appendix A

R Code for Kalman Filter Mean Reverting Model

This section summarizes *R* software code for a state space model via Kalman Filter Mean Reverting (KFMR) model used in this thesis corresponding to the following lines:

Lines 1-102	: Kalman Filter and Smoother algorithm code
Lines 103-154	: KFMR in Chapter 4 code
Lines 155-206	: TvLMM in Chapter 5 code
Lines 207-382	: Model C in Chapter 6 code

Note that Kalman Filter and Smoother algorithm code used in this thesis is a modified version of that described in Shumway and Stoffer (2006).

```
1 ######################################################################
2 ########## Kalman Filter and Smoother Algorithm   ####################
3 ######################################################################
4 ################### Kalman Filter ####################################
5 ######################################################################
6 kfilter=function(num,Y,A,mu0,Sigma0,Phi,Kappa,Alpham,Q,H){
7  Phi=as.matrix(Phi)                    # Phi
8  pdim=nrow(Phi)
9  Y=as.matrix(Y)                        # Y_t
10 qdim=ncol(Y)
11 B<-as.matrix(A)                       # A_t
12 Alphap=array(NA, dim=c(pdim,1,num))   # Alphap=Alpha_t^{t-1}  Prediction Alpha
```

```
 Pp=array(NA, dim=c(pdim,pdim,num))   # Pp=P_t^{t-1}  Prediction Alpha Variance
 Alphaf=array(NA, dim=c(pdim,1,num))  # Alphaf=Alpha_t^t  Filter Alpha
 Pf=array(NA, dim=c(pdim,pdim,num))      # Pf=p_t^t  Filter Alpha Variance
 v=array(NA, dim=c(qdim,1,num))          # Innovation
 sig=array(NA, dim=c(qdim,qdim,num))     # Innovation Variance
 K=array(NA, dim=c(pdim,qdim,num))       # Kalman gain
 Alpham=array(Alpham, dim=c(pdim,1,num))   # Alphamean
 Kappa=as.matrix(Kappa,nrow=pdim,ncol=1)   # Kappa
#########################################################
############### Kalman Filter t=1  #####################
#########################################################
# initial values
 Alpha00=as.matrix(mu0, nrow=pdim, ncol=1)             # Mu_0
 P00=as.matrix(Sigma0, nrow=pdim, ncol=pdim)           # Sigma_0
#################### Prediction ########################
 Alphap[,,1]=Phi%*%(Alpha00-Alpham[,,1]) + Alpham[,,1] # Prediction Alpha
 Pp[,,1]=Phi%*%P00%*%t(Phi)+Q                       # Prediction Alpha variance
#################### Filtering  ########################
 A=diag(B[1],nrow=qdim,ncol=qdim)                      # A_1
 sigtemp=A%*%Pp[,,1]%*%t(A)+H                          # Innovation variance
 sig[,,1]=(t(sigtemp)+sigtemp)/2                       # Symmetric matrix
 siginv=solve(sig[,,1])                             # Innovation variance inverse
 K[,,1]=Pp[,,1]%*%t(A)%*%siginv                        # Kalman gain
 v[,,1]=Y[1,]-A%*%(Alphap[,,1])-Kappa                  # Innovation
 Alphaf[,,1]=Alphap[,,1]+K[,,1]%*%v[,,1]               # Filter Alpha
 Pf[,,1]=Pp[,,1]-K[,,1]%*%A%*%Pp[,,1]                  # Filter Alpha variance
####################  MLE ##############################
 sigmat=as.matrix(sig[,,1], nrow=qdim, ncol=qdim)         # Innovation variance
 like =-0.5*(log(det(sigmat))+t(v[,,1])%*%siginv%*%v[,,1])# Loglikelihood Func.
########################################################
##########  Filter  iterations  t=2,...,n ##############
########################################################
 for (i in 2:num){
##################### Prediction  #######################
 Alphap[,,i]=Phi%*%(Alphaf[,,i-1]-Alpham[,,i])+ Alpham[,,i]
 Pp[,,i]=Phi%*%Pf[,,i-1]%*%t(Phi)+Q
##################### Filtering  ########################
 A=diag(B[i],nrow=qdim,ncol=qdim)
 sigtemp=A%*%Pp[,,i]%*%t(A)+H
 sig[,,i]=(t(sigtemp)+sigtemp)/2
 siginv=solve(sig[,,i])
 K[,,i]=Pp[,,i]%*%t(A)%*%siginv
 v[,,i]=Y[i,]-A%*%(Alphap[,,i])-Kappa
 Alphaf[,,i]=Alphap[,,i]+K[,,i]%*%v[,,i]
 Pf[,,i]=Pp[,,i]-K[,,i]%*%A%*%Pp[,,i]
###################### MLE ##############################
 sigmat=as.matrix(sig[,,i], nrow=qdim, ncol=qdim)
 like=like-(0.5*(log(det(sigmat))+t(v[,,i])%*%siginv%*%v[,,i]))
```

```
 }
 like=like-(((num*qdim)/2)*log(2*pi))
 list(Alphap=Alphap,Pp=Pp,Alphaf=Alphaf,Pf=Pf,like=like,Innov=v,sig=sig,Kn=K)
}
########################################################
############## Kalman Smoother #########################
########################################################
ksmooth=function(num,Y,A,mu0,Sigma0,Phi,Kappa,Alpham,Q,H){
 kf=kfilter(num,Y,A,mu0,Sigma0,Phi,Kappa,Alpham,Q,H)
 pdim=nrow(as.matrix(Phi))
 Alphas=array(NA, dim=c(pdim,1,num))         #Smoothing Alpha_n
 Ps=array(NA, dim=c(pdim,pdim,num))          #Smoothing Alpha_n variance
 J=array(NA, dim=c(pdim,pdim,num))
########################################################
##########  Smoothing inital condition t=n   ############
########################################################
# Smoothing Alpha_n = Filter Alpha_n
# Smoothing Alpha_n Variance= Filter Alpha_n Variance
########################################################
 Alphas[,,num]=kf$Alphaf[,,num]
 Ps[,,num]=kf$Pf[,,num]
########################################################
##########  Smoothing iterations t=n,...,2 ##############
########################################################
 for(k in num:2)  {
 J[,,k-1]=(kf$Pf[,,k-1]%*%t(Phi))%*%solve(kf$Pp[,,k])
 Alphas[,,k-1]=kf$Alphaf[,,k-1]+J[,,k-1]%*%(Alphas[,,k]-kf$Alphap[,,k])
 Ps[,,k-1]=kf$Pf[,,k-1]+J[,,k-1]%*%(Ps[,,k]-kf$Pp[,,k])%*%t(J[,,k-1])
 }
########################################################
##########  Smoothing  iteration t=1  ##################
########################################################
# initial values
 Alpha00=mu0
 P00=Sigma0
########################################################
 J0=as.matrix((P00%*%t(Phi))%*%solve(kf$Pp[,,1]), nrow=pdim, ncol=pdim)
 Alpha0n=as.matrix(Alpha00+J0%*%(Alphas[,,1]-kf$Alphap[,,1]), nrow=pdim, ncol=1)
 P0n= P00 + J0%*%(Ps[,,k]-kf$Pp[,,k])%*%t(J0)
 list(Alphas=Alphas,Ps=Ps,Alpha0n=Alpha0n,P0n=P0n,J0=J0,J=J,Alphap=kf$Alphap,Pp=kf$Pp,Alphaf=kf$Alphaf,Pf=kf$Pf,Innov=kf$Innov,sig=kf$sig,like=kf$like,Kn=kf$K)
}

####################################################################
########################### Chapter 4    ##########################
####################################################################
library(numDeriv)
```

```r
set.seed(1013314)
##########################################################
########################  Read Data  #####################
##########################################################
sp<-read.csv("P1ISE.csv", header=T)
sp<-data.frame(sp)
m<-length(sp[,1])
############# R_it - R_ft and R_mt - R_ft #######
st<-matrix(NA,m-1,20)                    # R_it
risk<-matrix(NA,m-1,1)                     # R_ft
risk<-(((1+(sp[2:m,22]/100))^(1/52))-1)  # Weekly R_ft
for(i in 2:21){                            # R_it=(log(P_t)-log(P_(t-1)))-R_ft
st[,i-1]<-diff(log(sp[,i]))-risk
}
st<-data.frame(st)                  # Data used in analyse
#########################################################
############### Kalman Filter and  Smoother  ############
#########################################################
# k=2,...,20 :> 19 Industry Sector Portfolios
rm<-st[,1]                          # R_mt - R_Ft  ISE Market
rt<-st[,k]                          # R_it - R_Ft  Sector Portfolio_i
# OLS estimates for initialization
datam<-data.frame(cbind(rt,rm))
fm<-lm(rt~rm-1,data=datam)
sumy<-summary(fm)
HH<-deviance(fm)/df.residual(fm)
coef<-cbind(sumy$coefficients[1,1],sumy$coefficients[1,2])
ols<-(coef[1,])
print(ols)
# MLE Process
# Use Transformed parameters, then convert back to unconstrained parameters
Linn=function(param){
 phi=param[1]^2/(1+(param[1]^2))    # Phi
 sigw=exp(param[2])                 # Q
 sigeps=exp(param[3])               # H
 kappa=0                            # Kappa
 betam=param[4]                     # Beta mean
 mu0=c(ols[1])                      # Mu_0
 sigma0<-c(ols[2])                  # Sigma_0
 sigma0<-sigma0%*%sigma0
 kf=kfilter(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)     # Kalman Filter
 return(-kf$like)
}
#Define the starting values for the unknown parameters.
#Use optim package to estimate the unknown parameters.
#Use the estimated parameters in the smoother process.
ksmooth(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)

```

```
####################################################################
#########################    Chapter 5    ###########################
#####################################################################
library(numDeriv)
set.seed(1013314)
##########################################################
########################  Read Data  #####################
##########################################################
sp<-read.csv("msci.csv", header=T)
sp<-data.frame(sp)
m<-length(sp[,1])
############# R_it - R_ft and R_mt - R_ft #######
st<-matrix(NA,m-1,19)                      # R_it
risk<-matrix(NA,m-1,1)                     # R_ft
risk<-(((1+(sp[2:m,21]/100))^(1/52))-1)   # Weekly R_ft
for(i in 2:20){                            # R_it=(log(P_t)-log(P_(t-1)))-R_ft
st[,i-1]<-diff(log(sp[,i]))-risk
}
st<-data.frame(st)                         # Data used in analyse
########################################################
############### Kalman Filter and  Smoother  ###########
########################################################
# k=2,...,19 :> 18 global markets
rm<-st[,1]                               # R_mt - R_ft Market
rt<-st[,k]                               # R_it - R_ft for Country_i
# OLS estimates for Initialization
datam<-data.frame(cbind(rt,rm))
fm<-lm(rt~rm,data=datam)
sumy<-summary(fm)
HH<-deviance(fm)/df.residual(fm)
coef<-cbind(sumy$coefficients[1:2,1],sumy$coefficients[1:2,2])
ols<-(coef[1:2,])
print(ols)
# MLE Process
# Use Transformed parameters, then convert back to unconstrained parameters
Linn=function(param){
 phi=param[1]^2/(1+(param[1]^2))       # Phi
 sigw=exp(param[2])                  # Q
 sigeps=exp(param[3])                  # H
 kappa=param[4]                      # Kappa
 betam=param[5]                      # Beta Mean
 mu0=c(ols[2,1])                     # Mu_0
 sigma0<-c(ols[2,2])                 # Sigma_0
 sigma0<-sigma0%*%sigma0
 kf=kfilter(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)  # Kalman Filter
 return(-kf$like)
}
#Define the starting values for the unknown parameters.
```

```
#Use optim package to estimate the unknown parameters.
#Use the estimated parameters in the smoother process.
ksmooth(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)

######################################################################
######################   Chapter 6    ###############################
############ Kalman Filter and  Smoother  #########################
######################################################################
# 2---10 Developed
# 11---19 Developed
rm<-st[,1]                  # R_mt - R_Ft Market
rt<-st[,2:10]               # R_it - R_Ft Country_i
# OLS estimates for Initialization
datam<-data.frame(cbind(rt,rm))
cols<-matrix(c(0,0),2,2)
HH<-rep(0,9)
k<-1
while (k<=9){
fm<-lm(datam[,k]~datam[,10],data=datam)
sumy<-summary(fm)
HH[k]<-deviance(fm)/df.residual(fm)
coef1<-cbind(sumy$coefficients[1,1],sumy$coefficients[1,2])
coef2<-cbind(sumy$coefficients[2,1],sumy$coefficients[2,2])
coef<-rbind(coef1,coef2)
sols<-(coef[1:2,])
cols<-rbind(cols,sols)
k<-k+1
}
ols<-cols[3:20,]
# MLE Process
# Use Transformed parameters, then convert back to unconstrained parameters
Linn=function(param){
 phi1=param[1]^2/(1+(param[1]^2))       # phi 1
 phi2=param[2]^2/(1+(param[2]^2))       # phi 2
 phi3=param[3]^2/(1+(param[3]^2))       # phi 3
 phi4=param[4]^2/(1+(param[4]^2))       # phi 4
 phi5=param[5]^2/(1+(param[5]^2))       # phi 5
 phi6=param[6]^2/(1+(param[6]^2))       # phi 6
 phi7=param[7]^2/(1+(param[7]^2))       # phi 7
 phi8=param[8]^2/(1+(param[8]^2))       # phi 8
 phi9=param[9]^2/(1+(param[9]^2))       # phi 9
 sigw1=abs(sqrt(exp(param[10])))        # Q 1
 sigw2=abs(sqrt(exp(param[11])))        # Q 2
 sigw3=abs(sqrt(exp(param[12])))        # Q 3
 sigw4=abs(sqrt(exp(param[13])))        # Q 4
 sigw5=abs(sqrt(exp(param[14])))        # Q 5
 sigw6=abs(sqrt(exp(param[15])))        # Q 6
 sigw7=abs(sqrt(exp(param[16])))        # Q 7
```

```
 sigw8=abs(sqrt(exp(param[17])))        # Q 8
 sigw9=abs(sqrt(exp(param[18])))        # Q 9
 corw1= param[19]/sqrt(1+(param[19]^2)) # Rho
 sigeps1=(exp(param[20]))              # H 1
 sigeps2=(exp(param[21]))              # H 2
 sigeps3=(exp(param[22]))              # H 3
 sigeps4=(exp(param[23]))              # H 4
 sigeps5=(exp(param[24]))              # H 5
 sigeps6=(exp(param[25]))              # H 6
 sigeps7=(exp(param[26]))              # H 7
 sigeps8=(exp(param[27]))              # H 8
 sigeps9=(exp(param[28]))              # H 9
 Kappa1=param[29]                      # Kappa 1
 Kappa2=param[30]                      # Kappa 2
 Kappa3=param[31]                      # Kappa 3
 Kappa4=param[32]                      # Kappa 4
 Kappa5=param[33]                      # Kappa 5
 Kappa6=param[34]                      # Kappa 6
 Kappa7=param[35]                      # Kappa 7
 Kappa8=param[36]                      # Kappa 8
 Kappa9=param[37]                      # Kappa 9
 Betam1=param[38]                      # Beta mean 1
 Betam2=param[39]                      # Beta mean 2
 Betam3=param[40]                      # Beta mean 3
 Betam4=param[41]                      # Beta mean 4
 Betam5=param[42]                      # Beta mean 5
 Betam6=param[43]                      # Beta mean 6
 Betam7=param[44]                      # Beta mean 7
 Betam8=param[45]                      # Beta mean 8
 Betam9=param[46]                      # Beta mean 9
# Phi Matrix
 phi=diag(c(phi1,phi2,phi3,phi4,phi5,phi6,phi7,phi8,phi9),9,9)
# Kappa and Beta mean vector
 kappa=c(Kappa1,Kappa2,Kappa3,Kappa4,Kappa5,Kappa6,Kappa7,Kappa8,Kappa9)
 betam=c(Betam1,Betam2,Betam3,Betam4,Betam5,Betam6,Betam7,Betam8,Betam9)
# H Matrix
 sigeps<-diag(c(sigeps1,sigeps2,sigeps3,sigeps4,sigeps5,sigeps6,sigeps7,sigeps8,
     sigeps9))
# Q Matrix
sigw<-matrix(0,9,9)
sigw[1,1] <- sigw1^2
sigw[1,2] <- corw1*sigw1*sigw2
sigw[1,3] <- corw1*sigw1*sigw3
sigw[1,4] <- corw1*sigw1*sigw4
sigw[1,5] <- corw1*sigw1*sigw5
sigw[1,6] <- corw1*sigw1*sigw6
sigw[1,7] <- corw1*sigw1*sigw7
sigw[1,8] <- corw1*sigw1*sigw8
```

```
sigw[1,9] <- corw1*sigw1*sigw9
sigw[2,1] <- corw1*sigw2*sigw1
sigw[2,2] <- sigw2^2
sigw[2,3] <- corw1*sigw2*sigw3
sigw[2,4] <- corw1*sigw2*sigw4
sigw[2,5] <- corw1*sigw2*sigw5
sigw[2,6] <- corw1*sigw2*sigw6
sigw[2,7] <- corw1*sigw2*sigw7
sigw[2,8] <- corw1*sigw2*sigw8
sigw[2,9] <- corw1*sigw2*sigw9
sigw[3,1] <- corw1*sigw3*sigw1
sigw[3,2] <- corw1*sigw3*sigw2
sigw[3,3] <- sigw3^2
sigw[3,4] <- corw1*sigw3*sigw4
sigw[3,5] <- corw1*sigw3*sigw5
sigw[3,6] <- corw1*sigw3*sigw6
sigw[3,7] <- corw1*sigw3*sigw7
sigw[3,8] <- corw1*sigw3*sigw8
sigw[3,9] <- corw1*sigw3*sigw9
sigw[4,1] <- corw1*sigw4*sigw1
sigw[4,2] <- corw1*sigw4*sigw2
sigw[4,3] <- corw1*sigw4*sigw3
sigw[4,4] <- sigw4^2
sigw[4,5] <- corw1*sigw4*sigw5
sigw[4,6] <- corw1*sigw4*sigw6
sigw[4,7] <- corw1*sigw4*sigw7
sigw[4,8] <- corw1*sigw4*sigw8
sigw[4,9] <- corw1*sigw4*sigw9
sigw[5,1] <- corw1*sigw5*sigw1
sigw[5,2] <- corw1*sigw5*sigw2
sigw[5,3] <- corw1*sigw5*sigw3
sigw[5,4] <- corw1*sigw5*sigw4
sigw[5,5] <- sigw5^2
sigw[5,6] <- corw1*sigw5*sigw6
sigw[5,7] <- corw1*sigw5*sigw7
sigw[5,8] <- corw1*sigw5*sigw8
sigw[5,9] <- corw1*sigw5*sigw9
sigw[6,1] <- corw1*sigw6*sigw1
sigw[6,2] <- corw1*sigw6*sigw2
sigw[6,3] <- corw1*sigw6*sigw3
sigw[6,4] <- corw1*sigw6*sigw4
sigw[6,5] <- corw1*sigw6*sigw5
sigw[6,6] <- sigw6^2
sigw[6,7] <- corw1*sigw6*sigw7
sigw[6,8] <- corw1*sigw6*sigw8
sigw[6,9] <- corw1*sigw6*sigw9
sigw[7,1] <- corw1*sigw7*sigw1
sigw[7,2] <- corw1*sigw7*sigw2
```

```
sigw[7,3] <- corw1*sigw7*sigw3
sigw[7,4] <- corw1*sigw7*sigw4
sigw[7,5] <- corw1*sigw7*sigw5
sigw[7,6] <- corw1*sigw7*sigw6
sigw[7,7] <- sigw7^2
sigw[7,8] <- corw1*sigw7*sigw8
sigw[7,9] <- corw1*sigw7*sigw9
sigw[8,1] <- corw1*sigw8*sigw1
sigw[8,2] <- corw1*sigw8*sigw2
sigw[8,3] <- corw1*sigw8*sigw3
sigw[8,4] <- corw1*sigw8*sigw4
sigw[8,5] <- corw1*sigw8*sigw5
sigw[8,6] <- corw1*sigw8*sigw6
sigw[8,7] <- corw1*sigw8*sigw7
sigw[8,8] <- sigw8^2
sigw[8,9] <- corw1*sigw8*sigw9
sigw[9,1] <- corw1*sigw9*sigw1
sigw[9,2] <- corw1*sigw9*sigw2
sigw[9,3] <- corw1*sigw9*sigw3
sigw[9,4] <- corw1*sigw9*sigw4
sigw[9,5] <- corw1*sigw9*sigw5
sigw[9,6] <- corw1*sigw9*sigw6
sigw[9,7] <- corw1*sigw9*sigw7
sigw[9,8] <- corw1*sigw9*sigw8
sigw[9,9] <- sigw9^2
# m_0 vector
mu0<-matrix(c(ols[2,1],ols[4,1],ols[6,1],ols[8,1],ols[10,1],ols[12,1],ols[14,1],
    ols[16,1],ols[18,1]),9,1)
#Sigma_0 matrix
sigma0<-diag(c(ols[2,2],ols[4,2],ols[6,2],ols[8,2],ols[10,2],ols[12,2],ols
    [14,2],ols[16,2],ols[18,2]),9,9)
sigma0<-sigma0%*%sigma0
kf=kfilter(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)     # Kalman Filter
return(-kf$like)
 }
#Define the starting values for the unknown parameters.
#Use optim package to estimate the unknown parameters.
#Use the estimated parameters in the smoother process.
ksmooth(n,rt,rm,mu0,sigma0,phi,kappa,betam,sigw,sigeps)
```

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.

Algieri, B. (2012). Price volatility, speculation and excessive speculation in commodity markets: Sheep or shepherd behaviour? *ZEF-Discussion Papers on Development Policy No. 166*, pages 1–40.

Anderson, B. and Moore, J. (1979). *Optimal Filtering*. Englewood Cliffs, Prentice Hall, NJ.

Barone-Adesi, G. (1985). Arbitrage equilibrium with skewed asset returns. *The Journal of Financial and Quantitative Analysis*, 20(3):299–313.

Bekaert, G. and Harvey, C. R. (1997). Emerging equity market volatility. NBER Working Papers 5307, National Bureau of Economic Research, Inc.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, 45(3).

Black, F. (1976). Studies of stock price volatility changes. In *Proceedings of the 1976 Meetings of the American Statistical Association*, Business and Economical Statistics Section, pages 177–181. Proceedings of the 1976 Meetings of the American Statistical Association.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics*, 72(3):498–505.

Bos, T. and Newbold, P. (1984). An empirical investigation of the possibility of stochastic systematic risk in the market model. *The Journal of Business*, 57(1):35–41.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7(3):265–296.

Brooks, R. D., Faff, R. W., and Lee, J. H. H. (1992). The form of time variation of systematic risk: Some Australian evidence. *Applied Financial Economics*, 2(4):191–198.

Brooks, R. D., Faff, R. W., and McKenzie, M. D. (1998). Time-varying beta risk of Australian industry portfolios: A comparison of modelling techniques. *Australian Journal of Management*, 23(1):1–22.

Brooks, R. D., Faff, R. W., and McKenzie, M. D. (2002). Time varying country risk: An assessment of alternative modelling techniques. *European Journal of Finance*, 8(3):249–274.

Campbell, J., Lo, A., and MacKinlay, A. (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ, USA.

Chang, T. and Nieh, C.-C. (2001). International transmission of stock price movements among Taiwan and its trading partners: Hong Kong, Japan and the United States. *Review of Pacific Basin Financial Markets and Policies*, 04(04):379–401.

Chiang, T. C., Jeon, B. N., and Li, H. (2007). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance*, 26(7):1206–1228.

Choudhry, T. and Wu, H. (2009). Forecasting ability of GARCH vs Kalman Filter method: evidence from daily UK time-varying beta. *The European Journal of Finance*, 15(4):437–444.

Christie-David, R. and Chaudhry, M. (2001). Coskewness and cokurtosis in futures markets. *Journal of Empirical Finance*, 8(1):55–81.

Christoffersen, P. (2003). *Elements of Financial Risk Management*. Elsevier Science.

Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik*, 31(4):377–403.

Cuthbertson, K. and Nitzsche, D. (2005). *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*. Financial Economics and Quantitative Analysis Series. Wiley.

Danielsson, J. (2011). *Financial Risk Forecasting: The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab*. The Wiley Finance Series. Wiley.

de Boor, C. (1978). *A Practical Guide to Splines*. Number 27 in Applied Mathematical Sciences. Springer-Verlag.

Durbin, J. and Koopman, S. (2001). *Time Series Analysis by State Space Methods*. Oxford Statistical Science Series. Clarendon Press.

Eilers, P. and Marx, B. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, 11(2):89–121.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory*, 11(01):122–150.

Fabozzi, F. J. and Francis, J. C. (1978). Beta as a random coefficient. *Journal of Financial and Quantitative Analysis*, 13(01):101–116.

Faff, R. W., Hillier, D., and Hillier, J. (2000). Time varying beta risk: An analysis of alternative modelling techniques. *Journal of Business Finance & Accounting*, 27(5-6):523–554.

Faff, R. W., Lee, J. H., and Fry, T. R. L. (1992). Time stationarity of systematic risk: Some Australian evidence. *Journal of Business Finance & Accounting*, 19(2):253–270.

Fahrmeir, L. and Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer Series in Statistics. Springer.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

Fang, H. and Lai, T.-Y. (1997). Co-kurtosis and capital asset pricing. *Financial Review*, 32(2):293–307.

Faraway, J. (2004). *Linear Models with R*. Chapman & Hall/CRC Texts in Statistical Science. Taylor & Francis.

Flavin, T. J., Hurley, M. J., and Rousseau, F. (2002). Explaining stock market correlation: A gravity model approach. *Manchester School*, 70:87–106.

Fletcher, R. (1987). *Practical Methods of Optimization; (2nd ed.)*. Wiley-Interscience, New York, NY, USA.

Friend, I. and Westerfield, R. (1980). Co-skewness and capital asset pricing. *The Journal of Finance*, 35(4):897–913.

Galagedera, D., Henry, D., and Silvapulle, P. (2002). Conditional relation between higher moments and stock returns: Evidence from Australian data. In *Proceedings from the Econometric Society Australian Meeting*. Queensland University of Technology, Brisbane, Australia.

Ghalanos, A. (2013). *rugarch: Univariate GARCH models*. R package version 1.2-7.

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5):1121–1152.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Staff Report 157, Federal Reserve Bank of Minneapolis.

Hansen, L. P. and Jagannathan, R. (1997). Assessing specification errors in stochastic discount factor models. *The Journal of Finance*, 52(2):557–590.

Harvey, A. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, A. C. and Shephard, N. (1993). Structural time series models. In *Econometrics*, volume 11 of *Handbook of Statistics*, pages 261–302. Elsevier.

Harvey, C. R. (1995). Predictable risk and returns in emerging markets. Working Paper 4621, National Bureau of Economic Research.

Harvey, C. R. and Siddique, A. (2000a). Conditional skewness in asset pricing tests. *Journal of Finance*, 55(3):1263–1296.

Harvey, C. R. and Siddique, A. (2000b). Time-varying conditional skewness and the market risk premium. *Research in Banking and Finance*, 1:27–58.

Hastie, T. and Tibshirani, R. (1990). *Generalized Additive Models*. Monographs on statistics and applied probability. Chapman and Hall.

Hosking, J. R. M. (1980). The multivariate portmanteau statistic. *Journal of American Statistical Association*, 75(371):602–607.

Hung, D. C. (2007). Return explanatory ability and predictability of non-linear market models. Working paper, DU, Durham.

Hung, D. C.-H., Shackleton, M., and Xu, X. (2004). CAPM, higher co-moment and factor models of UK stock returns. *Journal of Business Finance & Accounting*, 31(1-2):87–112.

Hurvich, C. M. and Tsai, C.-l. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307.

Hutchinson, C. E. (1984). The Kalman Filter applied to Aerospace and Electronic systems. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-20(4):500–504.

Hwang, S. and Satchell, S. E. (1999). Modelling emerging market risk premia using higher moments. *International Journal of Finance & Economics*, 4(4):271–296.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Javid, A. Y. (2009). Test of higher moment capital asset pricing model in case of Pakistani equity market. Mpra paper, University Library of Munich, Germany.

Jayasuriya, S. A. (2011). Stock market correlations between China and its emerging market neighbors. *Emerging Markets Review*, 12(4):418–431.

Jondeau, E. and Rockinger, M. (2004). Optimal portfolio allocation under higher moments. Working papers 108, Banque de France.

Jong, P. D. (1991). The diffuse Kalman Filter. *The Annals of Statistics*, 19(2):1073–1083.

Jurczenko, E. and Maillet, B. (2006). The four-moment capital asset pricing model: Between asset pricing and asset allocation. In Jurczenko, E. and Maillet, B., editors, *Multi-moment asset allocation and pricing models*, chapter 6, pages 113–142. John Wiley & Sons.

Kalman, R. and Bucy, R. (1961). New results in linear filtering and prediction theory. *Transactions of the ASME–Journal of Basic Engineering*, 83:95–107.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45.

Kim, D. (1993). The extent of nonstationarity of beta. *Review of Quantitative Finance and Accounting*, 3(2):241–254.

Kohn, R. and Ansley, C. F. (1989). A fast algorithm for signal extraction, influence and cross-validation in state space models. *Biometrika*, 76(1):65–79.

Koopman, S. J. (1993). Disturbance smoother for state space models. *Biometrika*, 80(1):117–126.

Kraus, A. and Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *The Journal of Finance*, 31(4):1085–1099.

Kutner, M., Neter, J., Nachtsheim, C., and Wasserman, W. (2005). *Applied Linear Statistical Models*. The McGraw-Hill/Irwin series in operations and decision sciences. McGraw-Hill Professional Publishing.

Levy, H. and Markowitz, H. M. (1979). Approximating expected utility by a function of mean and variance. *American Economic Review*, 69(3):308–317.

Levy, H. and Sarnat, M. (1994). *Capital Investment and Financial Decisions*, chapter Appendix 12A, pages 359–361. Prentice Hall, 5 edition.

Lim, K. G. (1989). A new test of the three-moment capital asset pricing model. *Journal of financial and Quantiative Analysis*, 24:205–216.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47(1):13–37.

Liow, K. H. and Chan, L. C. W. J. (2005). Co-skewness and co-kurtosis in global real estate securities. *Journal of Property Research*, 22(2-3):163–203.

Liu, Y., Pan, M.-S., and Shieh, J. (1998). International transmission of stock price movements: Evidence from the U.S. and five Asian-Pacific markets. *Journal of Economics and Finance*, 22(1):59–69.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Mackinlay, A. C. and Richardson, M. P. (1991). Using generalized method of moments to test mean-variance efficiency. *The Journal of Finance*, 46(2):511–527.

Mahdi, E. (2011). *Diagnostic Checking, Time Series and Regression.* PhD thesis, The School of Graduate and Postdoctoral Studies, The University of Western Ontario, Ontario, Canada.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Meinhold, R. J. and Singpurwalla, N. D. (1989). Robustification of Kalman Filter models. *Journal of the American Statistical Association*, 84(406):479–486.

Mergner, S. (2009). *Applications of State Space Model in Finance.* Universitätsverlag Gőttingen.

Mergner, S. and Bulla, J. (2008). Time-varying beta risk of pan-European industry portfolios: A comparison of alternative modeling techniques. *European Journal of Finance*, 14(8):771–802.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, 34(4):768–783.

Nelson, D. B. and Cao, C. Q. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, 10(2):229–235.

Odabasi, A. (2003). Some estimation issues on betas: A preliminary investigation in the Istanbul Stock Exchange. *Bogazici Journal: Review of Social, Economic and Administrative Studies*, 17(2):1–11.

Pennacchi, G. (2008). *Theory of Asset Pricing.* The Addison-Wesley Series in Finance. ADDISON WESLEY Publishing Company Incorporated.

Peters, J. (2001). Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (skewed) student-t densities. Working paper, EAA Business School, University of Liege.

Petris, G. (2014). [R] Kalman Filter estimation@MISC. `https://stat.ethz.ch/pipermail/r-help/2007-November/145898.html`.

Petris, G., Petrone, S., and Campagnoli, P. (2009). *Dynamic Linear Models with R*. Use R! Springer.

Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R*. Springer, 1 edition.

Pulley, L. B. (1981). A general mean-variance approximation to expected utility for short holding period. *Journal of Financial and Quantitative Analysis*, 16(3):361–373.

Ranaldo, A. and Favre, L. (2005). Hedge fund performance and higher-moment market models. *The Journal of Alternative Investments*, 8(3):37–51.

Roberts, G. (1986). *Switched-capacitor State-space Filters*. .

Roll, R. (1977). A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of Financial Economics*, 4(2):129–176.

Rosenberg, B. (1973). Random coefficients models: The analysis of a cross-section of time series by stochastically convergent parameter regression. *Annuals of Economic and Social Measurement*, 2(4):397–428.

Ruiz, E. (2007-2008). Financial econometrics. Technical report, London School of Economics.

Saleem, K. (2009). International linkage of the Russian market and the Russian financial crisis: A multivariate {GARCH} analysis. *Research in International Business and Finance*, 23(3):243 – 256.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49.

Schaefer, S., Brealey, R., Hodges, S., and Thomas, H. (1975). Alternative models of systematic risk. In Elton, E. and Gruber, M., editors, *International Capital Markets: An Inter and Intra Country Analysis*, pages 150–161. ., Amsterdam, North Holland.

Schwarz, G. E. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Schweppe, F. (1965). Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*, 11(1):61–70.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.

Shumway, R. and Stoffer, D. (2006). *Time Series Analysis and Its Applications: With R Examples*. Springer texts in statistics. Springer.

Simaan, Y. (1993). Portfolio selection and asset pricing-three parameter framework. *Management Science*, 39(5):568–577.

Sunder, S. (1980). Stationarity of market risk: Random coefficients tests for individual stocks. *Journal of Finance*, 35(4):883–896.

Takatoshi, I. and Wen-Ling, L. (1993). Price volatility and volume spillovers between the Tokyo and New York stock markets. Working Paper 4592, National Bureau of Economic Research.

Tepmony, S. (2010). Detecting premium portfolios in higher-order moments CAPM. Master's thesis, Ca' Foscari University of Venice.

Tsay, R. (2005). *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics. Wiley.

Wang, P. and Moore, T. (2008). Stock market integration for the transition economics: Time-varying conditional correlation approach. *The Manchester School*, 76:116–133.

Welch, G. and Bishop, G. (2006). An introduction to the Kalman Filter. Technical report, University of North Carolina at Chapel Hill., Chapel Hill, NC.

Wells, C. (1994). Variable betas on the Stockholm exchange 1971-1989. *Applied Financial Economics*, 4(1):75–92.

Wells, C. (1996). *The Kalman Filter in Finance*. Kluwer Academic Publishers.

Wood, S. (2006). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC Texts in Statistical Science Series. Chapman & Hall/CRC.

Yavas, B. F. (2007). Benefits of international portfolio diversification. *Graziadio Business Review*, 10(2).

Ziemann, V. (2004). *Conditional Asset Pricing with Higher Moments*. PhD thesis, Humboldt-Universität zu Berlin.

Zivot, E. and Wang, J. (2006). *Modeling Financial Time Series with S-PLUS*. Springer New York.